University of Sussex

Radiative feedback effects during reionization

David Sullivan

Submitted for the degree of Doctor of Philosophy

University of Sussex

September 2017

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

Parts of this thesis have been undertaken in collaboration with other researches. Where this is the case, we acknowledge their work at the beginning of the relevant chapter.

Signature:

David Sullivan

UNIVERSITY OF SUSSEX

David Sullivan, Doctor of Philosophy

Radiative feedback effects during reionization

Summary

During the first billion years after the big bang, the large-scale cosmic web of structures we see today began to form. This was followed by the first stars and galaxies, which brought an end to the Dark Ages. These first luminous sources are thought to be the prime candidates which fuelled cosmic reionization, the last major phase transition of the Universe, from a neutral inter-galactic medium following recombination to the ionized state it remains in today. The physical processes which drive reionization encapsulate several areas of research, from cosmology and galaxy formation to radiative transfer and atomic physics. Even with the wealth of present-day observational information at our disposal, the processes are still not fully understood. Therefore we cannot model reionization analytically, instead turning to numerical simulations using observations to constrain our models.

We perform a suite of fully-coupled radiation-hydrodynamical simulations of galaxy formation in cosmological volumes to probe the self-feedback of galaxies during the Epoch of Reionization. This research focuses on the transport of gas from the intergalactic medium onto dark matter halos, and consequences for semi-analytical models of galaxy formation. To improve on existing methods, which constrain the halo baryon fraction during reionization, we develop and train an artificial neural network to predict this quantity based on the physical properties of haloes. We demonstrate that this model is independent of redshift and reionization history, and can be trivially incorporated into semi-analytical models of galaxy formation. We further probe the physical processes which allow ionizing photons to escape from galaxies to reionize the Universe, specifically how stellar evolution uncertainties such as binary populations influence this process. Finally, we investigate to what extent a relative supersonic drift velocity between baryons and dark matter, present at recombination, may suppress the formation of the first objects and fundamentally alter their evolution. To do this, we develop a new method based on cosmological zoom simulations to include this effect in boxes much larger than the coherence length of the relative velocity for the first time.

Acknowledgements

I would like to thank my supervisor, Ilian Iliev, for his guidance and support throughout this PhD, and for securing the vast amount of computing time we used over the years, without which this work would not have been possible. I'd like to pay special thanks to Joakim Rosdahl and Keri Dixon; your support and guidance came at a crucial time in my PhD, and I can comfortably say that my work would have suffered greatly without your help. I would also like to thank my second supervisor, Peter Thomas, for his extremely useful support and advice at times when I needed it most. There are many other people who have helped with my research, to name but a few; Paul Shapiro, Anson D'Aloisio, Junhwan Choi, Anastasia Fialkov, and Pierre Ocvirk. I found our discussions both helpful and enjoyable.

One man who I owe a debt of gratitude to is my A-level Physics and Maths tutor, Mr Milligan. You are one of the kindest people I have ever met, and your patience and dedication to my education me will never be forgotten. Without your help, I wouldn't have made it to where I am today.

During my time in Sussex, I have been incredibly lucky to have met so many great people. It will perhaps come as a shock to nobody that I start with my first office mate, Scott Clay. Sharing an office and house with you was the single most challenging aspect of my entire PhD, and despite your best efforts, I have actually made it through to the other side. I'd also like to make special mention of Aled Anderson; you joined me behind enemy lines for every Six Nations, and all of our Thusday night curry dates. There a few people in this world that I can say, with certainty, will remain a friend through the test of time. You both make this list.

Special mention also goes out to my other housemates; Craig Perl and David Treanor, my tea buddy; Benoît Fournier, who ensured I never worked too hard, my 4C15 office mates; Alexander Eggemeier, Hannah Ross, Chaichalit Boyd Srisawat, Azizah Hosein, and Antonio Vazquez, who put up with me for several years and provided office snacks, my first year office neighbours; William Watson and Aurel Schneider, who were always on hand to help, and last, but not least, Dániel Molnár, Steven Duivenvoorden, and Ridwan

Barbhuiyan. While this list is certainly not exhaustive, you have all helped to keep me sane over the past four years.

I would also like to thank my parents, John and Karen Sullivan, and my extended family for their undying support during my University life. Mum, Dad, I couldn't have asked for anything more from you. Thank you for all that you have done. Thanks also goes out to June Davies, for all your support and kindness over the years.

Finally, I would like to thank my girlfriend, Emma Board-Davies. Moving to Brighton and leaving you behind was the single most difficult decision I have ever had to make, but you have been with me every step of the way. I now look forward to us finally being able to begin our lives together.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 An Introduction to Cosmology

On large scales, the Universe is observed to be expanding. Galaxies separate from one another in the expanding flow, with recessional velocities proportional to their distance. This was first observed by Hubble and Humason (1931) and famously expressed by the well known Hubble law. If we were to reverse the motion of galaxies, then we conclude that all matter originated from a single point, with infinite density and extremely high temperature. This instant in time is known as the Big Bang, before which our understanding of the Universe and the laws of physics which govern it break down. Perhaps the strongest evidence for the Big Bang Theory came with the discovery of the Cosmic Microwave Background (CMB) by Penzias and Wilson (1965), almost entirely by accident. This remnant radiation was produced when the Universe cooled sufficiently, such that neutral atoms became energetically favourable, leading to the (re)combination of protons and electrons to form neutral hydrogen and helium. This was immediately followed by the decoupling of photons from matter and a transition from an opaque to transparent Universe, when the CMB was last scattered. This period is known as cosmological recombination and ushered in the "dark ages" (Rees, 1999).

Observations of the CMB, such as those by the recent Planck mission, reveal that the early Universe was remarkably homogeneous to roughly one part in 10^5. This is reflected in the temperature anisotropies observed on the sky, as shown in figure 1.1 courtesy of the Planck collaboration, which traces the underlying matter distribution. Over time, these small density perturbations grow under the influence of gravity, eventually coalescing hierarchically to form the first non-linear structures. The concordance model of cosmology, the ΛCDM model, describes the Universe as mainly consisting of dark energy (Λ) and cold

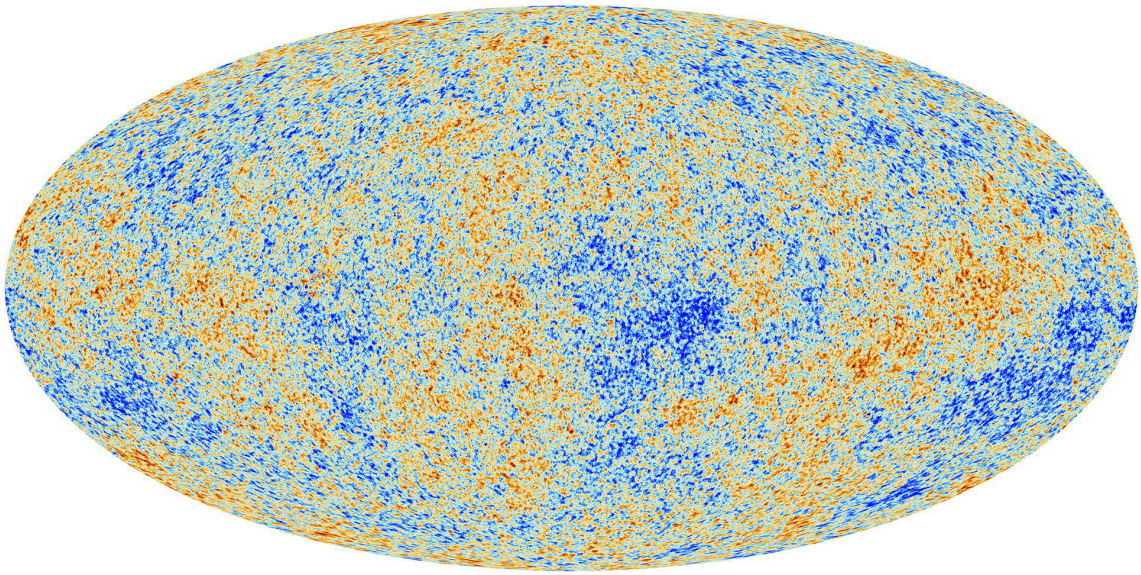

Figure 1.1: Mollweide projection of the CMB anisotropies on the sky as seen by the Planck satellite. Blue and red regions show cold and hot spots respectively, with a maximum deviation of approximately one part in 10^5. *Credit: ESA Planck Collaboration.*

dark matter (CDM). The exact properties of CDM remain a mystery, however evidence for its existence was found as early as the 1930s, when Zwicky (1933) studied the motions of galaxies and found that a significant fraction of matter in the Universe is unobserved and interacts only through gravity.

The notion of an invisible matter component, which went on to be known as "dark matter", can first be found in the literature as early as Poincare (1906). In ΛCDM, a model of kinematically cold weakly interacting dark matter is favoured. Naturally CDM particles quickly drop out of thermal equilibrium, as they do not undergo Thomson scattering with photons, therefore the first gravitationally bound structures are thought to be comprised of entirely dark matter. The resulting potential wells allow baryons to condense out of the intergalactic medium (IGM), paving the way for the formation of the first stars and galaxies which are thought to have started the process of reheating, thereby ushering in an end to the Dark Ages and the dawn of Reionization.

Many open questions surrounding reionization remain, such as when did it occur? What were the sources of reionization? Was it an extended process or instantaneous? What effect, if any, did it have on the first stars and galaxies? Very few observational probes are able to extend so far back in cosmic time, and even those which can, do not shed much light on this crucial period of galaxy formation. However, this is set to

change with the current generation of radio interferometers; such as LOFAR[1] and MWA[2] which are already in operation, the next generation such as SKA[3], and finally space based observatories such as JWST[4]. These observatories will each rely on theoretical studies of reionization to assist in understanding the wealth of data they are set to produce, with SKA alone set to output a staggering 160 terrabytes of data per second in its initial phase. To put this into context, after just one week the SKA will have produced more data than humans have managed to accumulate in our entire history.

Numerical modelling of this period is extremely challenging due to the huge range of relevant scales which must be accurately followed; from interactions of individual atoms to the formation of stars and galaxies. Furthermore, if we are to truly understand which sources drive reionization then we must also capture large volumes representative of the entire Universe, whilst simultaneously maintaining high enough resolution to resolve individual galaxies. The primary aim of this thesis is to understand the back reaction on the first stars and galaxies during cosmic reionization, using extremely high resolution, fully-coupled radiation-hydrodynamics in a cosmological context.

In this introductory chapter, we present an overview of our standard cosmological model and lay out the ground work for studying galaxy formation and the Epoch of Reionization (EoR). In section 1.2, the set of parameters used to describe our concordance model, and how they can be constrained from observations of the CMB, are described. This is followed by a brief overview of structure formation and linear theory, in section 1.3. Sections 1.4 and 1.5 provide a review of recent literature encapsulating the dark ages and the EoR, with a strong focus on feedback effects during this key episode of galaxy formation; which is further discussed in section 1.6. Finally, we conclude with a brief overview of this thesis in section 1.7.

1.2 Modern Cosmology

The General Theory of Relativity (GR) developed by Einstein in 1915 and the observations of Hubble and Humason (1931) which followed, mark the theoretical and observational pillars of modern cosmology. Hubble's law is the first observational confirmation of the *Cosmological principle*, which postulates that the Universe is both homogeneous and isotropic, i.e uniform in space at any given time and in all directions from any given point, on

[1] http://www.lofar.org
[2] http://www.mwatelescope.org
[3] http://skatelescope.org
[4] https://www.jwst.nasa.gov

large scales. The philosophical assumption behind this is that our place in the Universe is in no way a special one, and Einstein was remarkably successful in his assertion of the Cosmological principle.

A homogeneous and isotropic Universe provides the simplest solution to Einstein's field equations, for which the proper metric is the Friedmann-Lematre-Robertson-Walker (FLRW) metric. Evaluating the Einstein equations with this metric yields an analytical solution, the Friedmann equation:

$$H^2(t) = \left[\frac{\dot{a}(t)}{a(t)}\right]^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} \tag{1.1}$$

where $H(t)$ is the Hubble parameter, $a(t)$ is the scale factor which describes the expansion of the Universe at any given time, normalised to unity today. Due to the expansion of the Universe, the wavelengths of photons travelling towards us are Doppler shifted to higher wavelengths. This effect is known as *cosmological redshift*, and is related to the scale factor by:

$$z = \left(\frac{1}{a(t)}\right) - 1. \tag{1.2}$$

When discussing periods of cosmic history at specific times, it is convenient to instead quote the cosmological redshift, as this is an observable quantity in the light of distant objects.

The Friedmann equation determines how the size of the Universe depends on its contents and curvature, where ρ denotes the total energy density and k the geometrical curvature. Assuming a flat geometry ($k = 0$), which is preferred by observations, the Friendmann equation yields the critical density:

$$\rho_c(a) = \frac{3H^2(a)}{8\pi G}. \tag{1.3}$$

Observational evidence suggests that the total energy density is comprised of matter, radiation and dark energy in the form of a cosmological constant. It is convenient to parametrise these components in terms of the fractional contribution to the total density, which we define as the *density parameters*:

$$\Omega_{m,0} = \frac{\rho_{m,0}}{\rho_{c,0}};\ \Omega_{r,0} = \frac{\rho_{r,0}}{\rho_{c,0}};\ \Omega_{\Lambda,0} = \frac{\rho_{\Lambda,0}}{\rho_{c,0}}, \tag{1.4}$$

where the subscripts denote matter (dark and baryonic), radiation and dark energy respectively and the index 0 refers to their present day values. It is also useful to define the dimensionless Hubble parameter, h, when quoting results with respect to an overall scale

factor, such that $H_0 = H(t = 0) = 100\ h$ km s^{-1} Mpc^{-1}. By doing so, quantities can be expressed in a form which is independent of the assumed value of h, allowing for a less painful comparison of results. For a geometrically flat cosmology, we therefore find the total energy density parameter $\Omega_0 = \Omega_{m,0} + \Omega_{r,0} + \Omega_{\Lambda,0} = 1$.

One of these parameters, $\Omega_{r,0}$, has been directly measured with extremely high accuracy by the Cosmic Microwave Background Explorer (COBE) experiment (Smoot et al., 1992), which marked a milestone in our understanding of the CMB. As radiation and matter were coupled via Thompson scattering and thus in thermal equilibrium prior to recombination, theory predicts that the CMB spectrum should carry the shape of a blackbody. The COBE satellite confirmed this with extraordinary accuracy, measuring the CMB to be an almost perfect blackbody with temperature $T = 2.728 \pm 0.002$ K which corresponds to a radiation energy density $\Omega_{r,0} = 2.47 \times 10^{-5}\ h^{-2}$.

The energy density of matter and dark energy can also be constrained via the CMB, by studying the *power spectrum* of temperature anisotropies on the sky. The power spectrum describes the amplitude of fluctuations at different angular scales, and its shape is determined by acoustic oscillations in the photon-baryon plasma prior to recombination. The resonant frequencies and amplitudes of these acoustic waves depend on the composition and curvature of the Universe, due to the tight coupling of photons to matter. The position and amplitude of the first acoustic peak was first measured by COBE, with subsequent peaks measured for the first time by the Balloon Observations Of Millimetric Extragalactic Radiation ANd Geophysics (BOOMERanG; Netterfield et al., 2002) and Millimeter Anisotropy eXperiment IMaging Array (MAXIMA; Hanany et al., 2000) telescopes, albeit limited to small patches on the sky. The first all sky surveys came with the space based Wilkinson Microwave Anisotropy Probe (WMAP; Hinshaw et al., 2003) and Planck satellites (Planck Collaboration et al., 2011), with the latter reaching angular resolutions of a few arcminutes on the sky, as shown in figure 1.2.

The measured power spectrum (error bars) shows extremely good agreement with the concordance ΛCDM cosmological model (shaded region), which remarkably consists of just six baseline parameters. These include the previously discussed density parameters for matter and dark energy, $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$ respectively, the present day Hubble parameter, H_0, the tilt of the primordial power spectrum (i.e power law or spectral index), n_s, the amplitude of density fluctuations, σ_8, and finally the integrated optical depth due to reionization, τ. The CMB is sensitive to the number density of free electrons during reionization due to Thompson scattering of photons, therefore can be used to place

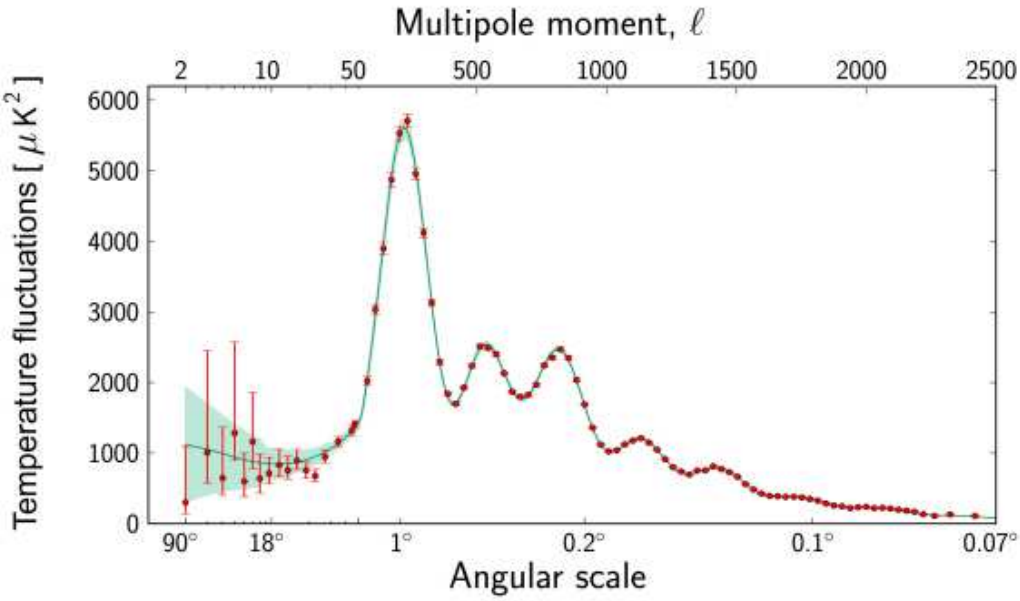

Figure 1.2: Angular power spectrum of CMB temperature anisotropies as observed by the Planck satellite. The green shaded region shows the best fit ΛCDM model. The exact positions and amplitudes of the peaks depend on the composition of the Universe, therefore can be used to put constraints on the cosmological density parameters. *Credit: ESA Planck Collaboration.*

Planck Collaboration 2015 Cosmological parameters		
Description	Symbol	Value
Spectral Scalar Index	n_s	0.9667 ± 0.0040
Hubble parameter	H_0	67.74 ± 0.46 km s^{-1} Mpc^{-1}
Matter density parameter	$\Omega_{m,0}$	0.3080 ± 0.0062
Dark energy density parameter	$\Omega_{\Lambda,0}$	0.6911 ± 0.0062
Amplitude of fluctuations at 8 Mpc h^{-1}	σ_8	0.8159 ± 0.0086
Optical depth due to reionization	τ	0.058 ± 0.012

Table 1.1: List of cosmological parameters from the Planck collaboration, with descriptions, symbols and values accurate as of the 2015 Planck release (Planck Collaboration et al., 2016b,a). The first five parameters are used to generate cosmological initial conditions from which we initialise our simulations, however their values may differ based on the assumed underlying cosmology. Where this is the case, their assumed values are given and briefly discussed.

constraints on τ and subsequently rule out certain models of reionization. The combination of WMAP and Planck along with multi-wavelength galaxy surveys has led to the era of precision cosmology, where we can determine the energy density parameters to the fourth decimal place. In table 1.1 we list the latest constraints on these parameters from the Planck 2015 data release and reionization constraints (Planck Collaboration et al., 2016b,a), for reference. Throughout this thesis these key parameters are used to generate cosmological initial conditions from which we simulate the EoR, however their values may change depending on the assumed underlying cosmological model. Where this is the case, the impact of varying these parameters is discussed in light of our findings.

The present day cosmological parameters show that the expansion of the Universe is currently dominated by dark energy, however, this has not always been the case. In figure 1.3, we illustrate the evolution of the energy densities of radiation, matter, and dark energy as a fraction of the critical density at fixed scale factor, $a(t)$. The continuity equation for a homogeneous mass density implies $\rho_m \propto a^{-3}$, i.e as the Universe expands and the scale factor increases, the matter mass density declines inversely with volume. The radiation energy density, which includes the CMB and three species of relativistic neutrinos, decreases as $\rho_r \propto a^{-4}$, as the density of photons is diluted as a^{-3} and the energy per photon declines as a^{-1} due to the redshifting of photon wavelength. Finally, assuming dark energy in the form of a cosmological constant, $\rho_\Lambda(a) = \rho_{\Lambda,0}$. The initial phase of expansion was dominated by radiation at all redshifts $z \gtrsim z_{eq} \simeq (\Omega_m/\Omega_r) - 1 \sim 2600$, after which matter domination begins. It is during this era that the CMB cools sufficiently to allow baryons to decouple from photons and begin to collapse, triggering cosmological recombination ($z_{\rm rec} \sim 1100$). Dark energy takes over as the primary driver of cosmic expansion by $z_\Lambda \simeq (\Omega_\Lambda/\Omega_m)^{1/3} - 1 \sim 0.4$, and remains dominant to this day.

Understanding how these components evolve over time is crucial if we are to set out a theoretical framework for the growth of structure. In the next section, we provide a brief statistical description of structure formation in the linear limit, when deviations from the cosmic mean density are still small, within the ΛCDM framework.

1.3 Structure Formation

In the previous section we described the evolution of the Universe as a whole, assuming homogeneity and isotropy on large scales. However, we know the Universe is inhomogeneous on small scales, as we see a vast array of structure from stars and planets to clusters of galaxies. In order to provide a connection between these two paradigms, we must

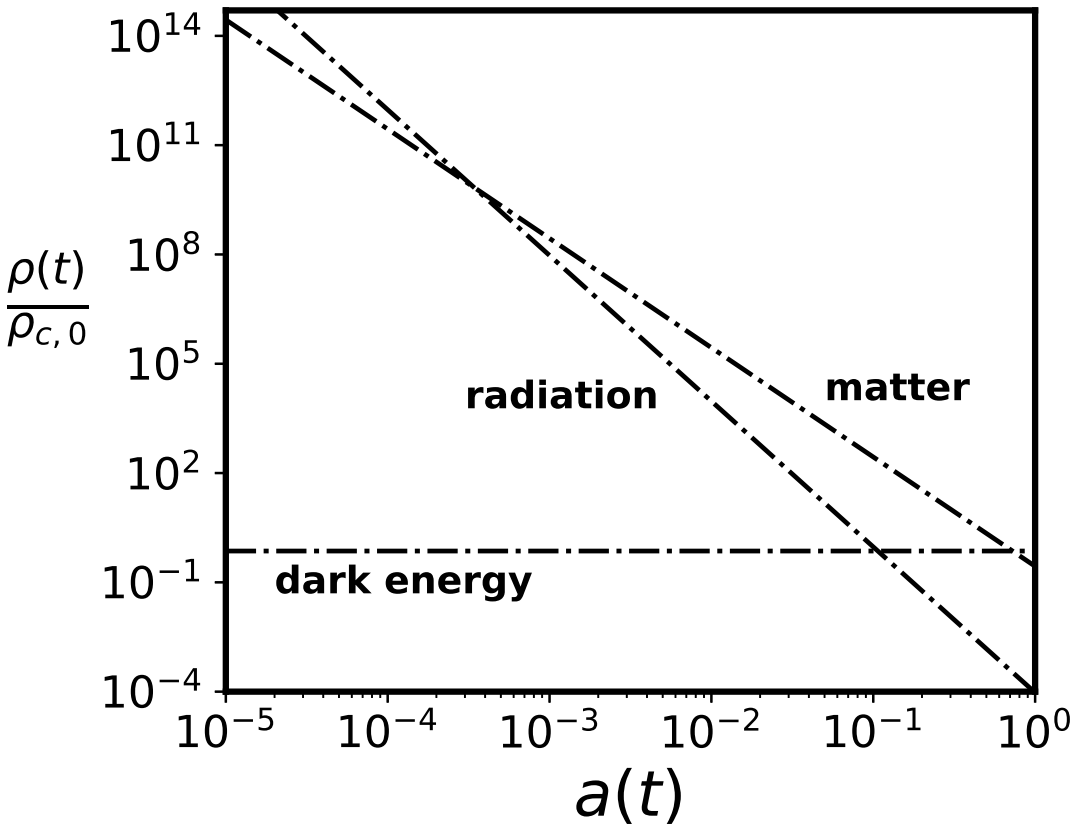

Figure 1.3: Evolution of the density of radiation, matter and dark energy with cosmic scale factor. The lines show the periods of cosmic time in which radiation, matter, and dark energy have dominated the cosmic expansion. Radiation matter equality marks the beginning of recombination, when photons decoupled and the CMB was last scattered. Matter domination proceeded during which the first stars and galaxies began to form. As the Universe continued to expand, the density of matter was diluted to the point where dark energy became dominant, where it remains today.

first understand the growth of initial perturbations in the density field from a theoretical standpoint. We are aided in this challenge by knowing that at the time of recombination, fluctuations in the density of matter were extremely uniform and therefore can be described using a perturbative approach. This is valid until the first non-linear structures begin to form, at which time we require a more thorough description. The three fundamental equations describing the conservation of mass and momentum, along with the gravitational potential can therefore be linearised to describe the initial growth of density perturbations. These equations are known as the *Euler* and *Poisson* equations of self-gravitating fluid-dynamics. In this section, we briefly outline these equations in their linearised form in order to introduce notations used later in this thesis. We show how starting with the resulting set of differential equations, gravitational instabilities arise which led to the formation of the first non-linear structures in the Universe. The objects, known as *dark matter haloes*, eventually become home to the first stars and galaxies.

1.3.1 Linear Theory

Let us consider some part of the Universe at the mean density $\bar{\rho}_m$, which we may define in terms of the fraction energy density of matter and the critical density at redshift z:

$$\bar{\rho}_m(z) = \Omega_{m,0}\rho_c(1+z)^3. \tag{1.5}$$

A convenient way to describe departures from the mean density is to introduce the dimensionless density contrast: $\delta(\mathbf{x}) = \rho(\mathbf{x})/\bar{\rho}_m - 1$, where $\mathbf{x}$ is the comoving spatial coordinate. Assuming initial perturbations are small, i.e $|\delta| \ll 1$, then a linear description is sufficient to describe their growth under gravity. The Euler equations of fluid dynamics in comoving coordinates, where the bulk velocity vanishes, can then be linearised as follows:

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla \cdot [(1+\delta)\mathbf{u}] = 0, \qquad \text{(continuity)} \tag{1.6}$$

$$\frac{\partial \mathbf{u}}{\partial t} + H\mathbf{u} + \frac{1}{a}(\mathbf{u}\cdot\nabla)\mathbf{u} = -\frac{1}{a}\nabla\phi - \frac{\nabla P}{a\bar{\rho}_m(1+\delta)}, \qquad \text{(Euler)} \tag{1.7}$$

$$\nabla^2\phi = 4\pi G\bar{\rho}_m a^2\delta, \qquad \text{(Poisson)} \tag{1.8}$$

where $\mathbf{u} \equiv \mathbf{v} - H\mathbf{r}$ is the peculiar velocity describing departures from the Hubble flow, $\mathbf{r}$ and $\mathbf{v}$ are the physical (proper) position and velocity of the fluid element, P is the pressure, and finally ϕ is the Newtonian gravitational potential. This set of equations requires closure in the form of an effective equation of state for the pressure term, determined by

thermodynamical processes. Generally, this can be written as a function of density and entropy, i.e $P = P(\rho, S)$ where S is the specific entropy. For an ideal non-relativistic monatomic gas, it can be shown that:

$$\frac{\nabla P}{\bar{\rho}_m} = \frac{1}{\bar{\rho}_m}\left[\left(\frac{\partial P}{\partial \rho}\right)_S \nabla\rho + \left(\frac{\partial P}{\partial S}\right)_\rho \nabla S\right], \tag{1.9}$$

$$= c_s^2 \nabla\delta + \frac{2}{3}(1+\delta)T\ \nabla S, \tag{1.10}$$

where $c_s = (\partial P/\partial\rho)_S$ is the adiabatic sound speed and T is the gas temperature. Using the above result and keeping terms that are first order in δ or $\mathbf{u}$ only, equations 1.6 - 1.8 can be combined to yield a second-order differential equation which describes the evolution of density perturbations with time:

$$\frac{\partial^2\delta}{\partial t^2} + 2H\frac{\partial\delta}{\partial t} = 4\pi G\bar{\rho}_m\delta + \frac{c_s^2}{a^2}\nabla^2\delta - \frac{2\bar{T}}{3a}\nabla^2 S, \tag{1.11}$$

where we have used the definition of the Hubble parameter in equation 1.1. For a pressure-less CDM fluid the final two terms on the right hand side vanish, while the second term on the left hand side is known as the Hubble drag term and tends to suppress the growth of density fluctuations due to the expanding Universe. For a pressure-less CDM fluid and spatially flat Universe, equation 1.11 has a general solution in the form of a growing and decaying mode:

$$\delta(\mathbf{x}, t) = A(\mathbf{x})D_+(t) + B(\mathbf{x})D_-(t), \tag{1.12}$$

where $D_+(t)$ is the growing mode, which dominates over the decaying mode $D_-(t)$. Substituting equation 1.12 into equation 1.11 for the growing mode only, and assuming a spatially flat FLRW Universe containing only matter and vacuum energy, the *linear growth factor* is given by (Heath, 1977; Peebles, 1980):

$$g(\Omega_{m,0}, \Omega_{\Lambda,0}) \equiv \frac{D_+(t)}{a} = \frac{5\Omega_{m,0}}{2}\int_0^1 \frac{da}{a^3\left(H(a)/H_0\right)^3}, \tag{1.13}$$

where $H(a)/H_0 = (\Omega_{m,0}a^{-3} + \Omega_{\Lambda,0})^{1/2}$ for a flat Universe. Therefore, until they become non-linear, density perturbations grow in amplitude in proportion to $D_+(t)$ while maintaining their shape in comoving coordinates.

1.3.2 Gravitational Instability

As density fluctuations evolve, they influence each other through gravity and therefore do not evolve independently of one another. However, if we instead consider these fluctuations as a stochastic superposition of plane waves with varying amplitude, then within the linear regime these waves *do* evolve independently of one another. The relationship between density waves $\delta_{\mathbf{k}}(t)$ and a spatial fluctuation $\delta(\mathbf{x}, t)$ at comoving position $\mathbf{x}$ and time t is then trivially given by the Fourier transform:

$$\delta_{\mathbf{k}}(t) = \frac{1}{(2\pi)^3} \int d^3\mathbf{x}\; e^{-i\mathbf{k}\cdot\mathbf{x}}\; \delta(\mathbf{x}, t), \tag{1.14}$$

where $\mathbf{k}$ is the *wave vector*, and is related to wavelength $\lambda = 2\pi/|\mathbf{k}|$. By taking the Fourier transform of equation 1.11 in this way, and using the fact that spacial derivatives are simplified in doing so as $\nabla = i\mathbf{k}$, we obtain:

$$\frac{d^2\delta_{\mathbf{k}}}{dt^2} + 2H\frac{d\delta_{\mathbf{k}}}{dt} = \left[4\pi G\bar{\rho}_m - \frac{k^2 c_s^2}{a^2}\right]\delta_{\mathbf{k}} - \frac{2}{3}\frac{\bar{T}}{a^2}k^2 S_{\mathbf{k}}, \tag{1.15}$$

where for simplicity we have ignored CDM, which in a thorough treatment would impact the growth of baryonic fluctuations via gravity, and hence coupled via the Poisson equation.

This equation gives the time evolution of density fluctuations at fixed spatial scales, where the density perturbation $\delta_{\mathbf{k}}$ and entropy perturbation $S_{\mathbf{k}}$ act as source terms. For a self-gravitating fluid in a non-expanding Universe, $H = 0$, where the initial perturbation is isentropic with adiabatic evolution, we can set $\mathbf{k}^2 S_{\mathbf{k}} = 0$ such that equation 1.15 yields the dispersion relation:

$$\frac{d^2\delta_{\mathbf{k}}}{dt^2} = -\omega^2\delta_{\mathbf{k}}, \quad \text{where} \quad \omega^2 = \frac{k^2 c_s^2}{a^2} - 4\pi G\bar{\rho}_m. \tag{1.16}$$

A stability criterion for gravitational collapse can then be derived by taking the special case of $\omega = 0$, which defines a characteristic length scale below which pressure waves can no longer support against collapse:

$$k_{\mathrm{J}} = \frac{2\pi a}{c_s}\sqrt{\frac{G\bar{\rho}_m}{\pi}}. \tag{1.17}$$

This scale is known as the *Jeans scale*, and has a corresponding proper length, $\lambda_{\mathrm{J}} = 2\pi/k_{\mathrm{J}}$, known as the *Jeans length*. For an expanding Universe, the Hubble drag term causes a small damping of acoustic waves with $\lambda < \lambda_{\mathrm{J}}$, along with a suppression in the growth of unstable waves with $\lambda > \lambda_{\mathrm{J}}$. By defining the corresponding mass scale as the mass enclosed within a sphere of radius $\lambda_{\mathrm{J}}/2$, we arrive at the *Jeans mass*:

$$M_{\rm J} = \frac{4\pi}{3}\bar{\rho}_m \left(\frac{\pi a}{k_{\rm J}}\right)^3, \tag{1.18}$$

This defines the minimum mass scale which is unstable to collapse, and depends on the adiabatic sound speed of the gas by equation 1.17. For example, prior to recombination the sound speed of the gas is relativistic, with $c_s = c/\sqrt{3}$, and the corresponding Jeans mass is[5] $M_{\rm J} \sim 10^{19}\ M_\odot$, where $M_\odot$ denotes one solar mass. Isentropic perturbations with scale sizes below this therefore cannot grow until after recombination, when the sound speed falls drastically, to of order the thermal velocities of neutral atoms, $c_s \sim 2 \times 10^{-5} c$. The Jeans mass decreases by roughly 10 orders of magnitude as a result, to $M_{\rm J} \sim 10^8\ M_\odot$. The Jeans mass continues to fall as the Universe expands and cools, allowing baryons to collapse and eventually go on to form the first stars and galaxies.

1.3.3 The Filtering Mass

Within linear theory, the Jeans mass is related to the growth of perturbations only at a fixed time, $a(t)$. As the Jeans mass varies in time with the local sound speed, the overall suppression of growth instead depends on the time averaged Jeans mass, known as the *filtering mass*. This mass scale can be thought of as the cosmological analogue of the Jeans mass, and takes the same form as equation 1.18 only the corresponding length scale is instead replaced with the comoving wavenumber k_F, which is the characteristic scale over which baryonic perturbations are smoothed as compared to the dark matter (Gnedin and Hui, 1998):

$$\frac{1}{k_F^2(t)} = \frac{1}{D_+(t)} \int_0^t dt' a^2(t) \frac{\ddot{D}_+(t') + 2H(t')\dot{D}_+(t')}{k_{\rm J}^2(t')} \int_{t'}^t \frac{dt''}{a^2(t'')}, \tag{1.19}$$

At high redshift ($z \gtrsim 2$, where $\Omega_m \to 1$), the scale factor scales roughly as a power law with time, $a \propto t^{2/3}$, and the growing mode $D_+(t)$ is proportional to a. Equation 1.19 therefore simplifies to:

$$\frac{1}{k_F^2(a)} = \frac{3}{a} \int_0^a \frac{da'}{k_{\rm J}^2(a')} \left(1 - \sqrt{\frac{a'}{a}}\right), \tag{1.20}$$

and the filtering mass is defined in the same way we defined the Jeans length:

$$M_F = \frac{4\pi}{3}\bar{\rho}_m \left(\frac{\pi a}{k_F}\right)^3. \tag{1.21}$$

[5] Note that the above analysis assumes a non-relativistic Newtonian fluid, which is not the case prior to recombination. However, this value is still accurate to roughly an order of magnitude.

A simple argument for the working with the filtering mass over the Jeans mass is given by Gnedin (2000b): the Jeans scale, k_J is related to the mean density at a single time only. The gas temperature can rise suddenly (and therefore also the virial temperature and the Jeans mass), however density may only vary over a dynamical time (approximately the Hubble time for linear perturbations). The filtering scale, k_F, on the other hand is only defined at a given time by considering the entire evolution of the Jeans scale up to that moment and hence accounting for the full thermal history of the gas. The evolution of the Jeans and filtering mass scales are plotted in figure 1.4, courtesy of Tseliakhovich et al. (2011), including the effects of a supersonic relative drift velocity between baryons and CDM which is discussed later in section 1.4. As the Universe cools, the Jeans mass falls due to the declining sound speed while the filtering mass remains roughly constant in the redshift range shown. It is initially much smaller than the Jeans mass as baryon fluctuations are suppressed after recombination and must therefore catch up to the CDM fluctuations.

1.3.4 The Power Spectrum

Choosing to work in the Fourier domain leads to a statistical description of the initial perturbations, therefore any comparison to observations will also have to be statistical. The characteristic amplitude of each **k**-mode describes the density fluctuation δ on some spatial scale λ, therefore by squaring the amplitudes of these modes, and averaging over the whole ensemble at each scale k, we obtain a statistic called the *power spectrum*:

$$P(\mathbf{k}) = \langle \delta_{\mathbf{k}} \delta^*_{\mathbf{k}} \rangle = \langle |\delta_{\mathbf{k}}|^2 \rangle = (2\pi)^3 \delta^D(\mathbf{k} - \mathbf{k}') P(\mathbf{k}), \quad \text{and} \quad \mathcal{P}(k) = \frac{k^3 P(k)}{2\pi^2}, \tag{1.22}$$

where $\mathcal{P}(k)$ is the dimensionless power spectrum[6], and is essentially the variance of amplitudes at any given k (where $k = |\mathbf{k}|$ for an isotropic Universe). Above, we have used the fact that one can set $\delta^*_{\mathbf{k}} \equiv \delta_{\mathbf{k}}$ from the relation between modes with opposite wavevectors for any real field, $\delta_{\mathbf{k}}(-\mathbf{k}) = \delta^*_{\mathbf{k}}(\mathbf{k})$ (Peacock, 1999). Species specific power spectra are computed using transfer functions, $T(k)$, which encode the physical process which modify the power spectrum during matter radiation equality when k modes enter the horizon, and evolved forward linearly from time $t1$ to time t:

$$P(k,t) = P_0(k) T^2(k) \frac{D_+^2(k,t)}{D_+^2(k,t1)}, \tag{1.23}$$

[6]Usually, the notation $\Delta^2(k)$ is reserved for the dimensionless power spectrum, however we adopt $\mathcal{P}(k)$ to avoid confusion with the overdensity, Δ.

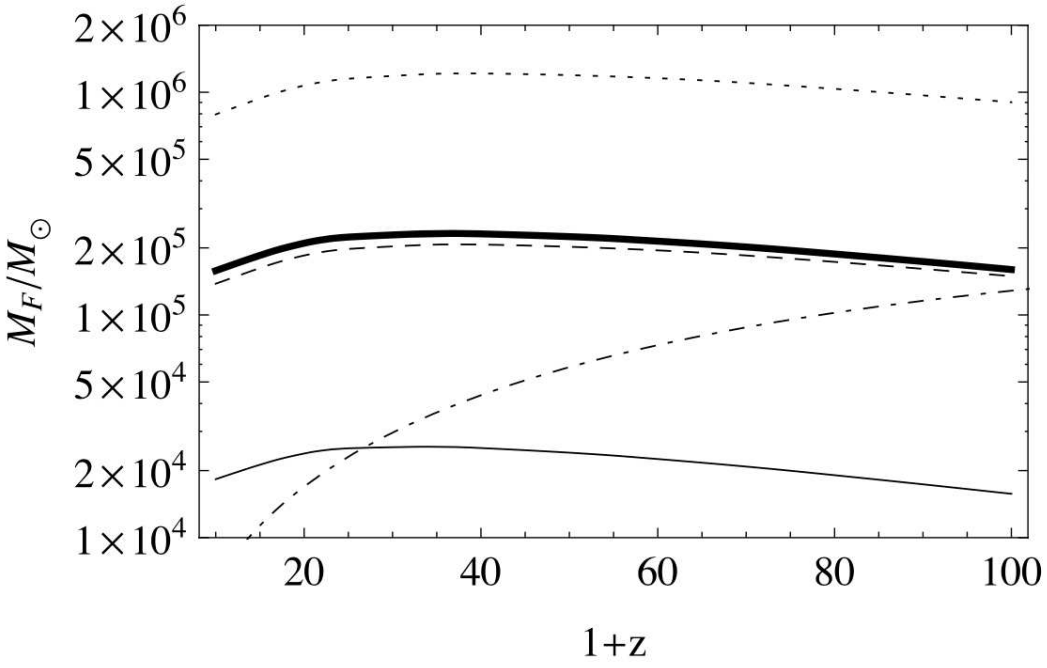

Figure 1.4: Evolution of the Jeans and Filtering masses with redshift, where the dot-dashed line shows the Jeans mass computed by equation 1.18 and the thin solid line shows the Filtering mass computed by equation 1.21. The dashed and dotted lines also show the evolution of the Filtering mass, but in regions where the relative velocity between baryons and CDM is $v_{\rm bc} = \langle v_{\rm bc}^2 \rangle^{1/2}$ and $v_{\rm bc} = 2\langle v_{\rm bc}^2 \rangle^{1/2}$, respectively, while the thick solid line shows the average over the entire velocity distribution (the relative velocity is discussed later in section 1.4). The relative velocity acts as an increased sound speed, which increases the Filtering mass and must be dissipated by the virialisation of the gas. *Credit:* Tseliakhovich et al. (2011).

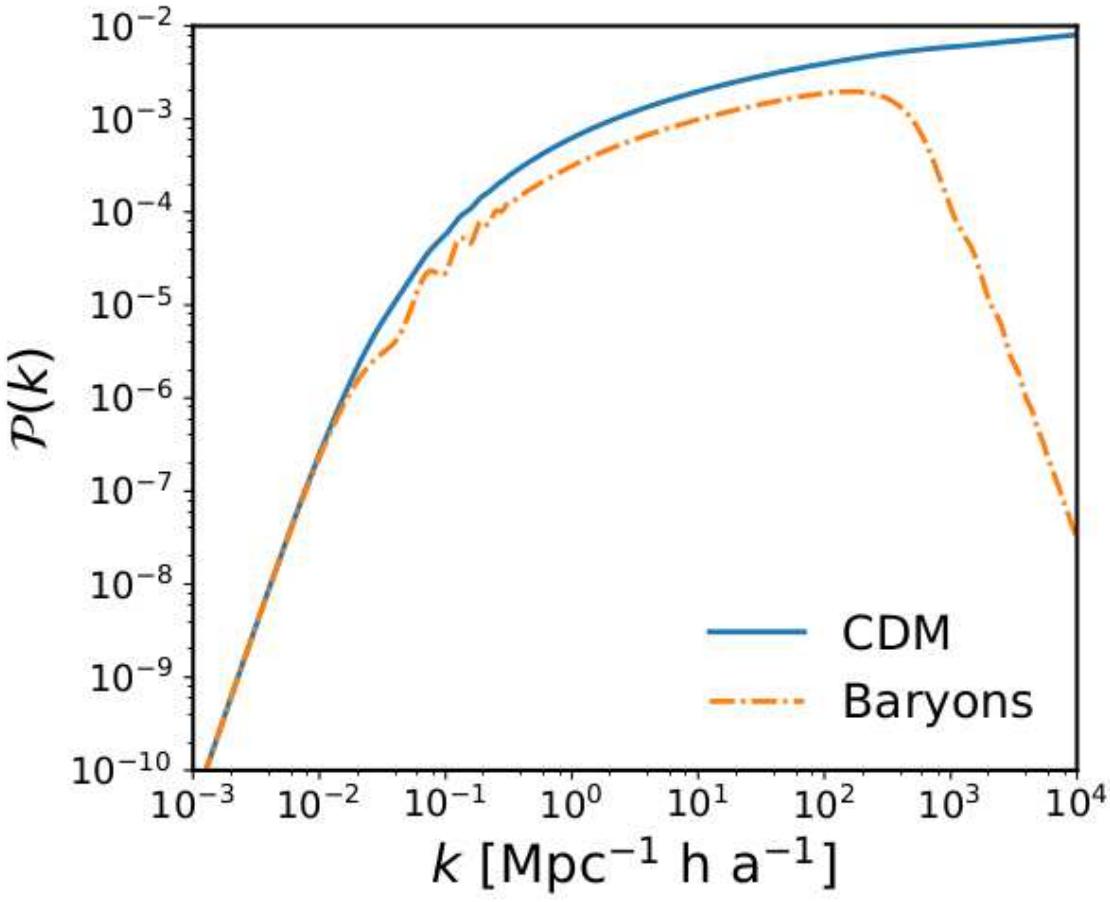

Figure 1.5: Dimensionless power spectra of CDM and baryons at redshift $z = 100$, computed using CAMB (Lewis et al., 2000). The power spectrum has the interesting property that it completely describes the statistics of any Guassian random field, therefore can to generate cosmological initial conditions at redshifts where the fluctuations are still linear.

where $P_0(k) = A_s k^{n_s}$ is the primordial (inflationary) power spectrum, A_s is the amplitude of scalar fluctuations, and n_s is the scalar index introduced in section 1.2. In figure 1.5, we show the dimensionless power spectra of CDM (solid; blue) and baryons (dot-dashed; orange) at $z = 100$, computed using the CAMB code (Lewis et al., 2000).

The features at $k \sim 10^{-1}$ Mpc^{-1} h a^{-1} in the baryon spectrum show the acoustic oscillations imprinted in the matter distribution prior to recombination, that led to acoustic peaks in the CMB. While less obvious, these features are also present in the power spectrum of CDM, as it is dragged by baryons via gravity. Baryonic power is suppressed on small scales, $k \gtrsim 3 \times 10^2$ Mpc^{-1} h a^{-1}, due to these pressure-induced oscillations, while CDM quickly fall out of thermal equilibrium following the big bang and therefore can collapse to small scales. The normalisation is set by the variance of the power spectrum:

$$\sigma^2 \equiv \langle \delta^2 \rangle = \int_0^\infty 4\pi k^2 P(k) dk, \tag{1.24}$$

which is usually smoothed on comoving scales of 8 h^{-1} Mpc, thus denoted σ_8 as listed in table 1.1.

1.3.5 Dark Matter Haloes and Spherical Collapse

While it is assumed that $M_{\rm J}$ and M_F accurately reflect the scales at which gravitational collapse occurs, they only describe whether an initial perturbation begins to grow. As the densities of dark matter and baryons (δ_c and δ_b) begin to grow above unity, they become non-linear and their density profiles start to evolve. At this point, dark matter shells may cross baryonic shells due to the different dynamics between the two fluids, hence the mass enclosed within a baryonic shell may increase with time. Eventually the collapsing dark matter drags baryons with it, causing collapse even below the filtering mass. The minimum mass of the resulting non-linear baryonic object is not well known, but for a spherically symmetric tophat density perturbation one can appeal to the spherical collapse model.

The exact details of this model are not important for this thesis, however it details the chain of events leading to collapse. Assuming an spherically symmetric perturbation in the (pressureless) CDM only, the perturbation initially expands with the Hubble flow but begins to slow due to gravity. It eventually breaks away from the Hubble flow and reaches a maximum radius $r_{\rm max}$ and time $t_{\rm max}$, before it turns around and begins to contract until it finally collapses under self-gravity. Linear theory estimates the turnaround and collapse overdensities to be $\delta_{\rm turnaround} = (3/20)(6\pi)^{2/3} = 1.06$ and $\delta_{\rm collapse} = (3/20)(12\pi)^{2/3} = 1.69$ (Mo et al., 2010). In fact for a perfectly spherically symmetric, pressureless model the overdensity at collapse reaches infinity. In practice, the collapse is halted long before this singularity due to a process known as "virialization". Matter within the overdensity therefore obeys the virial theorem, which relates the kinetic and potential energies of the system such that $2K + U = 0$, where K and U denote the kinetic and potential energies respectively, and describes the global properties of any dynamical system. For a flat Universe with $\Omega_m + \Omega_\Lambda = 1$, and assuming that virialization occurs at $t_{\rm vir} = 2t_{\rm max}$, the virial overdensity can be approximated as:

$$\Delta_{\rm vir} \equiv \rho(\mathbf{x})/\bar{\rho}_m \approx (18\pi^2 + 82x - 39x^2)/\Omega_m(t_{\rm vir}), \tag{1.25}$$

where $x = \Omega_m(t_{\rm vir}) - 1$ (Bryan and Norman, 1998). During matter domination, when

$\Omega_m = 1$, this simplifies to $\Delta_{\rm vir} \approx 18\pi^2 \approx 178$, therefore a collapsed halo will always have an average density within it that is roughly 200 times the cosmic mean density at the epoch of collapse. These collapsed dark matter spheres are known as *dark matter haloes*, for which we define physical properties based on the radius of a sphere with a mean enclosed density $\Delta_{\rm vir}$ times larger than the mean. The virial mass, for example, is defined as the mass enclosed within this sphere and with mean density $\Delta_{\rm vir}\bar{\rho}_m$, such that $M_{\rm vir} = (4\pi/3)\, R_{\rm vir}^3 \Delta_{\rm vir}\bar{\rho}_m$. We use this definition to detect haloes in our simulations throughout this thesis.

1.4 The Dark Ages

The thermal history of the IGM during the dark ages has important implications for the first stars and galaxies. Reheating is thought to have occurred as a result of the first luminous objects, however weak structure formation shocks post-recombination could result in significant entropy generation in the kinetically-cold IGM (see, for example; Furlanetto and Loeb, 2004). Linear perturbation treatments, accounting for all relevant particle species, show that a small but supersonic relative velocity existed in most parts of the Universe between baryons and CDM (Hu and Sugiyama (1995), Ma and Bertschinger (1995)). At the epoch of recombination, baryons are moving with a root-mean-square (RMS) velocity of 30 km s^{-1}, which was coherent over comoving Mpc scales (Tseliakhovich and Hirata (2010)). This coincides with a drop in the local sound speed from relativistic ($\sim c/\sqrt{3}$) to the thermal velocities of neutral atoms ($\sim 2{\times}10^{-5}$c), therefore baryons stream supersonically through the CDM with mach numbers up to $\mathcal{M} \simeq 5$ at $z \sim 1100$. By this time CDM structures have already began to form, since linear perturbations grow once their density is greater than that of the primordial radiation. Since the drift velocity, v_{bc}, is a vector perturbation, it decays with time scaling as $(1+z)$, therefore is most important at high-redshifts (Fialkov, 2014). By $z = 50$, the typical streaming velocity is ~ 1.5 km s^{-1}, corresponding to an equivalent gas temperature $T_{bc} \sim m_p v_{bc}^2/k_{\rm B} = 270$ K$[(1+z)/50]^2$, which is well above the mean IGM temperature. Hence, the relative velocity acts as an increased (effective) sound speed (since it needs to be dispersed by the virialisation of the gas; Loeb and Furlanetto, 2013).

Recent studies have incorporated this effect into numerical simulations to probe the effect of shock heating on the first haloes. These require an extremely high dynamical range, capable of resolving the horizon and Jeans scale simultaneously. This is currently not possible, therefore several works have chosen to focus on either large or small scales

only (Greif et al., 2011; Maio et al., 2011; Stacy et al., 2011; Naoz et al., 2012, 2013). All have incorporated this relative velocity by simply adding bulk velocities to the gas in their initial conditions, in accordance with the analytical predictions of Tseliakhovich and Hirata (2010). Maio et al. (2011) performed three simulations with initial bulk velocities of 0, 30 and 60 km s^{-1} at the recombination epoch and scaled to their initial simulation redshift of $z = 100$. Larger bulk velocities were found to delay the onset of population III star formation by approximately 10^7 years and suppress the gas fraction of the first haloes by upto 50%. Reionization is also therefore delayed by approximately the same time and initially driven by massive objects, as dwarf haloes cannot retain their gas and form galaxies. Boosting of the bulk velocities in this way however is a crude approximation, neglecting much of the non-linear impact on structure formation at redshifts larger than the initial redshift of the simulation. O'Leary and McQuinn (2012) recently showed this approximation significantly under-estimates the impact on the linear growth factor.

They were the first to initialize, self-consistently using linear theory, gas and dark matter on scales where pressure has a non-negligible impact on the growth rate of perturbations and baryons do not trace dark matter. Gas is found downwind of dark matter haloes, and the maximum gas density can be suppressed by up to an order of magnitude for haloes in the mass range $10^4 M_\odot - 10^6 M_\odot$, therefore suppressing star formation in dwarf haloes. Dynamical friction acts to erase the relative velocity, decelerating the gas into the dark matter rest frame by $z = 20$, whilst also generating Mach cones and entropy around dark matter haloes. These shocks could potentially reheat the IGM, suggesting it was not kinetically cold at the time when baryons decelerated and condensed into dark matter haloes[7]. It is generally believed that at low mass, dwarf haloes drive reionization as massive haloes are rare (predicted from the halo mass function). The suppression of both the gas fraction and star formation rate of these dwarf haloes by this supersonic drift velocity therefore has important implications for reionization.

1.5 The Epoch of Reionization

Reionization was the second of two major phase transitions of the Universe, from a neutral IGM following recombination to the ionized state it remains today[8]. In its simplest form, this process can be described by three key stages across a redshift range of $\Delta z \sim 10$. During

[7] In the absence of shocks, the gas cools adiabatically, as $(1 + z)^2$, reaching 10K at $z = 20$.

[8] Primordial helium experiences the same phase transitions, with complete reionization (i.e HeI and HeII) occurring much later in the history of the Universe than hydrogen reionization. As hydrogen is far more abundant than helium, we usually refer to reionization as the reionization of hydrogen only.

the initial "pre-overlap" stage, the Universe is predominantly neutral except for small isolated H II regions surrounding individual stellar sources and dense neutral gas. The resulting ionization fronts expand into the low-density diffuse gas in the IGM, eventually overlapping and hence defining the "overlap" stage. This occurs in approximately 10% of the Hubble time, leading to a sharp peak in the ionizing background. Finally, remaining neutral regions in dense clumps of gas are ionized, completing the process (Barkana and Loeb, 2001; Loeb and Barkana, 2001; Bromm and Larson, 2004; Ciardi and Ferrara, 2005; Choudhury and Ferrara, 2006; Furlanetto et al., 2006).

The exact definition of "complete" is somewhat subjective, however in this thesis we consider the Universe to be completely ionized when the volume weighted neutral fraction falls to $\sim 10^{-3}$. These assumptions suggest that reionization proceeds outside-in, with ionized regions expanding quickly into voids before finally reionizing the dense neutral regions. Initial simulations of reionization (Gnedin and Ostriker, 1997; Gnedin, 2000a; Ciardi et al., 2000; Ricotti et al., 2002), which were restricted to small volumes (< 10 Mpc on a side), widely supported an outside-in completion. The first large-scale simulations, spanning a few hundred Mpc on a side and performed by Iliev et al. (2006b), showed that in fact reionization progresses inside-out. In this scenario, high-density regions undergo reionization prior to voids due to the strong exponential growth of the source population which can only be captured by including the additional large-scale density fluctuations. Several later works came to similar conclusions, consistent with an inside-out completion (see for example, McQuinn et al., 2007; Trac and Cen, 2007). Investigations by Iliev et al. (2014) found that the additional large-scale density fluctuations, not present in small volume realizations, acts to strongly modulate the local number of ionizing sources. This additional power also introduces additional patchiness in the ionization field, which is not statistically captured in small volumes.

Simulations play an important role in the qualitative understanding of observations. Our primary interest here is the first stars and (proto)galaxies, for which direct observations of their formation are not yet possible. The Ly-α forest in the spectra of high redshift quasars, motivated the early theoretical works of Ikeuchi (1986) and Rees (1986) into the impact of a UV background on the assembly of low mass structures. The latter argued that dark matter haloes with circular virial velocities $v_c \sim 30\mathrm{km\ s^{-1}}$ are capable of retaining gas during UV photoheating (i.e maintaining the cosmic mean baryon fraction $\bar{f}_b = \Omega_b/\Omega_m \simeq 0.2$). Their assessment assumes that the characteristic mass (M_c), at which a halo retains half its gas mass, $\bar{M}_b = 0.5 f_b M_{tot}$, corresponds to the Jeans mass with virial

temperature $T_{vir} \sim 10^3 - 10^4$K, i.e the typical temperatures of photoionized gas during reionization (below which potential wells are too small). The characterization of suppression during reionization by a single mass scale is desirable for semi-analytical models of dwarf galaxy formation, due to its simplicity and ability to be easily implemented into existing models. Initial attempts to quantify M_c using semi-analytical models (SAMs) (Babul and Rees, 1992; Efstathiou, 1992) found agreement on the notion of low mass halo suppression, however concluded that the Jeans mass overpredicted M_c by roughly an order of magnitude. Later, Shapiro et al. (1994) and Gnedin (2000b) proposed that the filtering mass was appropriate, and found strong agreement with 3D cosmological hydrodynamics simulations (Quinn et al., 1996; Weinberg et al., 1997).

Modern cosmological hydrodynamics simulations have re-visited this problem to explore the accuracy of the fit put forward by Gnedin (2000b), commonly used in SAMs of galaxy formation both pre and post reionization. In general, the filtering mass provides a good fit to the minimum halo mass that is able to retain gas during reionization. However, the same is not true for the minimum halo mass that is able to accrete gas, once a quasi-homogeneous UV background is present (Hoeft et al., 2006; Okamoto et al., 2008; Naoz et al., 2009). For cold gas coupled to the Hubble flow at $z \sim 20$, the filtering mass sets the scale above which gas is able to fragment into smaller parcels and the Jeans mass scales as $\rho^{1/2}$. These parcels are unable to fragment further as they collapse into dark matter haloes adiabatically and decouple from the Hubble flow, therefore one expects the filtering scale to characterize these masses well. On the other hand, diffuse hot gas in the IGM after reionization has a temperature floor of $\sim 10^4$K and hence cannot collapse adiabatically. Instead the collapse is isothermal and the Jeans mass scales as $\rho^{-1/2}$, therefore gas fragments during collapse (as density increases). This motivated alternative analytical predictions by Hoeft et al. (2006) and Okamoto et al. (2008), which instead characterize M_c by the virial temperature of the gas evaluated at the halo virial density (opposed to the mean density as done previously). They find that a halo which can retain its gas after reionization can be up to an order of magnitude lower than the filtering mass. It should however be noted that none of the above models provide us with a complete picture of the physics of gas accretion during this period. In fact, Noh and McQuinn (2014) followed the trajectories of gas particles in the $T - n_H$ plane and found the above still overestimate the magnitude of suppression following reionization. Furthermore, they showed that spurious structure formation shocks during the adiabatic collapse phase act to further increase the mass threshold of accretion. The above studies also neglect a full

radiative transfer treatment coupled to the hydrodynamics, instead opting for an instantaneous quasi-homogeneous UV background at some predetermined redshift, e.g the UV background model of Haardt and Madau (2001).

Observational evidence for reionization is sparse, with results to date only able to constrain particular details of this epoch. The thermal state of the IGM can be probed via the Lyman-α (Ly-α) absorption line from high redshift quasars, due to the presence of neutral hydrogen. The large cross-section for Ly-α absorption ($\sigma([1+z]\nu_0)$; $\lambda_\alpha = 1215.67$ Å) leads to a Gunn-Peterson trough blue-ward of the Ly-α line for a completely neutral IGM. As a result, quasars only accurately probe the end of reionization as strong absorption features are present for neutral fractions as small as $x_{\rm HI} \equiv n_{\rm HI}/n_H \gtrsim 10^{-4}$ (Fan et al., 2006). Several quasars have been observed to exhibit a Gunn-Peterson trough at $z > 6.1$ (see figure 8 of Fan et al., 2006), which suggests reionization ended around this time. A further constraint on the redshift of reionization, $z_{\rm reion}$, arises from observations of the CMB polarization signal. Polarization is generated on scales $\ell \lesssim 20$ corresponding to the horizon scale at the EoR, due to Thompson scattering by free-electrons produced by reionization (Zaldarriaga, 1997; Hu and Sugiyama, 1995). The latest Planck forecasts place the optical depth due to reionization at $\tau = 0.058 \pm 0.012$, which corresponds to a reionization redshift of $z_{\rm reion} = 8.8 \pm 0.9$ (redshift-symmetric; Planck Collaboration et al., 2016a). The most promising observations, however, are expected to come from the next generation of radio telescopes aimed at observing the 21cm transition of atomic hydrogen. This transition is extremely weak and therefore, unlike the Ly-α Gunn-Peterson trough, does not saturate in the neutral IGM with an effective optical depth of 1% (Fan et al., 2006). Furthermore, the low-energy/frequency means it is sensitive to small temperature fluctuations in the IGM, and can therefore be observed against the CMB. Below, we briefly describe key observational constraints on the EoR, while directing the reader to Zaroubi (2013) for a more comprehensive review.

1.5.1 Observational Constraints

Gunn-Peterson Trough

High-redshift quasars, or quasi stellar objects (QSOs), probe the ionization state of the IGM via the absorption of Ly-α photons by residual neutral gas along the line of sight. This effect imprints a series of absorption lines in the spectra of QSOs, producing a feature known as the Ly-α forest. Gunn and Peterson (1965) predicted that given enough neutral gas along the line of sight, emission lines would be completely suppressed. This

became known as the Gunn-Peterson trough, for which the optical depth for a uniformly distributed IGM is given by:

$$\tau_{\rm GP}(z) = \frac{1.8\times10^5}{h}\Omega_0^{1/2}\left(\frac{\Omega_{b,0}h^2}{0.02}\right)\left(\frac{1+z}{7}\right)^{3/2}\left(\frac{n_{\rm HI}}{n_{\rm H}}\right), \tag{1.26}$$

where $\Omega_{b,0}$ is the energy density of baryons, $n_{\rm H}$ is the number density of hydrogen, and HI denotes neutral hydrogen. One caveat of trying to constrain reionization in this way is that even neutral fractions of $(n_{\rm HI}/n_{\rm H}) = 10^{-3}$ can result in $\tau \gg 1$. This effect was first observed in the spectra of a $z = 6.28$ quasar by Becker et al. (2001), leading to the conclusion that there must be some neutral gas present at this time. Subsequent studies of high-redshift QSOs in the Sloan Digital Sky Survey (SDSS) (e.g Fan et al., 2001, 2006) suggest that reionization was completed by $z \sim 6$, while the first direct upper limits on the mean neutral fraction by McGreer et al. (2011) constrain the mean neutral fraction to be $\bar{x}_{\rm HI} \lesssim 0.2$ at $z \lesssim 5.5$ and rising to $\bar{x}_{\rm HI} < 0.8$ at $z = 6.1$. They use the simple and robust statistic of the covering fraction of dark pixels in the Ly-α/β forests of high-redshift QSOs, and do not rely on assumptions about QSO continua, IGM density, the morphology of H II regions or ionizing background fields. Due to the high sensitivity of $\tau_{\rm GP}$ on the neutral fraction, as shown in equation 1.26, it is difficult to put anything but upper constraints on the redshift of overlap using this method.

Thompson Scattering Optical Depth of CMB Photons

As previously discussed, the CMB is sensitive to reionization due to Thompson scattering of photons off free electrons. Many authors have considered this effect in the context of the resulting suppression in the temperature anisotropies of the CMB (Peebles and Yu, 1970; Sunyaev and Zeldovich, 1972; Bond and Efstathiou, 1984; Ma and Bertschinger, 1995; Hu and White, 1997; Aghanim et al., 2008). As CMB photons can scatter off free-electrons along the entire line of sight from the observer to the surface of last scattering, the optical depth is given by the following integral over redshift:

$$\tau = \int_0^{z_{\rm rec}} \sigma_{\rm T} n_e \frac{cH_0^{-1}dz}{(1+z)\sqrt{\Omega_{m,0}(1+z)^3+\Omega_{\Lambda,0}}}, \tag{1.27}$$

where σ_T is the Thompson scattering cross-section, and n_e is the number density of electrons. It is important to note that as τ is an integrated quantity, it is not particularly sensitive to varying models for the redshift evolution of n_e. Furthermore, constrains on $z_{\rm reion}$ using this value assume an instantaneous reionization, which is a highly-unlikely scenario as most models predict an extended process driven by stars and QSOs (Zaroubi, 2013).

The latest Planck forecasts place the optical depth due to reionization at $\tau = 0.058 \pm 0.012$, which corresponds to an instantaneous reionization redshift of $z_{\rm reion} = 8.8 \pm 0.9$ (redshift-symmetric; Planck Collaboration et al., 2016a).

Hyperfine 21-cm Transition of Neutral Hydrogen

The most promising observations are expected to come from the next generation of radio telescopes aimed at observing the redshifted 21-cm transition line of atomic hydrogen. It originates due to a transition between the two hyperfine levels of the hydrogen $1s$ ground state. In the hydrogen atom, electrons and protons both have spin 1/2, corresponding to the two allowed values of the quantum spin parameter $\pm 1/2$. The spins can either be aligned in parallel or antiparallel, where the former corresponds to slightly higher energy. The energy difference between these two spin states corresponds to 5.9×10^{-6} eV, with frequency $\nu_{21} = 1420.4057$ MHz and wavelength $\lambda_{21} = 21.1061$ cm. The probability for transition between the two states is given by the Einstein $A_{10} = 2.85 \times 10^{-15}\ s^{-1}$ coefficient, therefore the emission of 21-cm radiation from neutral hydrogen atoms is extremely rare. While it cannot be observed in a laboratory on Earth due to collisional de-excitation processes, the large abundance of neutral hydrogen in the pre-reionization IGM means it is possible to probe the dark ages with radio observatories. We briefly summarise 21-cm radiation as a probe of the dark ages and reionization here, but direct the reader to Furlanetto et al. (2006); Pritchard and Loeb (2012); Furlanetto (2016) for a more comprehensive review.

In radio astronomy, the measured quantity is the differential brightness temperature, $\delta T_b = T_b - T_{\rm CMB}$, which measures deviations in the brightness temperature of the IGM, T_b, from the CMB. The emission or absorption of 21-cm photons from neutral hydrogen is governed by the spin temperature, T_s, which is defined through the equation:

$$\frac{n_1}{n_0} = 3e^{-T_*/T_s}, \tag{1.28}$$

where n_1 and n_0 are the relative number densities of atoms in the triplet and singlet states of the hyperfine level respectively, and $T_* = 0.0681$ K is the temperature corresponding to the 21-cm wavelength. The spin flip transition quickly reaches equilibrium with the CMB, coupling T_s and $T_{\rm CMB}$. The differential brightness temperature can then be defined as a function of the spin temperature (Mellema et al., 2006b):

$$\delta T_b = \frac{T_s - T_{\rm CMB}}{1+z}(1 - e^{-\tau_{21}}), \tag{1.29}$$

where τ_{21} is the optical depth of a cloud of hydrogen, and $e^{-\tau_{21}}$ is the transmission probability.

The local Ly-α flux couples the spin temperature to the local kinetic temperature of hydrogen through the Wouthuysen-Field effect. This coupling plays a role in producing a difference in the temperature of neutral hydrogen and the cosmic microwave background at the end of the Dark Ages and the beginning of the epoch of reionization, allowing for detailed 21cm studies at during the EoR. Simulations will prove essential to the qualitative understanding of observations from the upcoming SKA observatory, such as the detailed modelling of Semelin et al. (2007); Semelin (2016); Ross et al. (2017). The SKA is expected to provide full 3D mapping of the 21-cm signal, from which astrophysical constraints can be extracted using inversion methods. Interestingly, Shimabukuro and Semelin (2017) demonstrated that artificial neural networks (ANNs) provide a promising alternative to inversion using the Markov Chain Monte Carlo method, as they require only a sparse exploration of the parameter space and a large volume of real world training data, which SKA will deliver.

1.6 Galaxy Formation and The First Stars

Haloes provide the basic conceptual unit for the deeply non-linear regime, and host the formation of galaxies via the cooling and fragmentation of baryons. While gravity controls the assembly through accretion and merging, radiative cooling controls gas supply through condensation. In White and Rees (1978), a two-stage model of galaxy formation was first proposed, where dark matter haloes are assumed to grow via mergers and accretion while luminous galaxies form at their centre from the cooling and fragmentation of gas. This bottom-up hierarchical model is characteristic of ΛCDM, as small structures are the first to form and dynamical relaxation effects erase any trace of progenitors in the merger process. While the processes involved in galaxy formation are complex are poorly understood, below we briefly discuss the essential cooling and heating mechanisms, as well as the assumed properties of the first stars and galaxies.

1.6.1 Heating and Cooling Mechanisms

As gas crosses the virial sphere, a strong virialization shock results in the kinetic infall energy being thermalized and therefore heated to of order the virial temperature, which we define as:

$$T_{\rm vir} = \frac{1}{2}\frac{\mu m_{\rm p}}{k_{\rm B}} V_{\rm c}^2, \tag{1.30}$$

where

$$V_{\rm c} = \left(\frac{G{\rm M}_{\rm vir}}{R_{\rm vir}}\right)^{1/2} \tag{1.31}$$

is the circular velocity of the halo at the virial radius, μ is the mean molecular weight, m_p is the proton mass, and $k_{\rm B}$ is the Boltzmann constant. The shock forms as the infall velocity of the gas, which is supersonic, vanishes at the halo centre due to gas pressure. As the resulting shock propagates out through the halo, gas which crosses the shock is heated. This increases the sound speed and makes the interior flow subsonic, resulting in the conversion of the infall kinetic energy to thermal energy. In order for this to persist, the gas internal to the shock must have pressure to drive the propagating shock fronts. Following shock heating, the gas proceeds to be supported in a quasi-static equilibrium, which fills the halo with hot gas at densities several orders of magnitude below typical interstellar medium (ISM) densities. Here, it cools and slowly contracts to a disc where it can eventually form stars provided the cooling time is shorter than the gas free-fall time, $t_{\rm ff} \sim \sqrt{1/G\rho}$ (Benson, 2010). The cooling process is therefore essential in the formation of central galaxies, and in determining important properties such as the star-formation rate.

In the temperature range 10^4 K $< T_{\rm vir} < 10^6$ K, or hot enough to collisionally excite hydrogen lines and partially or fully ionize the gas, atomic cooling dominates via a number of excitation and de-excitation processes, including recombination radiation and collisional excitation with subsequent decay. Above these temperatures, the gas is fully collisionally ionized and cools mainly via free-free emission (i.e.., Bremsstrahlung). Haloes whose virial temperatures are below the atomically cooling limit, i.e. $T_{\rm vir} \sim 8000$ K, are unable to collisionally ionize hydrogen and, therefore, are expected to be entirely neutral in absence of ionizing radiation. If gas is metal-free, the dominant cooling process in such haloes is via the rotational and/or vibrational lines of molecular hydrogen (Abel et al., 1997) and is crucial for the formation of the first stars. Finally, at high redshifts ($z > 6$), inverse Compton scattering of cosmic microwave background photons also provides an efficient cooling mechanism (Compton cooling). We show cooling functions for primordial ($Z = 0$) and metal enriched gas in figure 1.6, courtesy of Agertz et al. (2013). The peaks in the cooling functions of primordial gas at temperatures $T \sim 15,000$ K and 10^5 K are due to the collisionally excited levels of H^0 and He^+ respectively, and are further enhanced in enriched gas due to the collisionally excited levels of heavier elements, such as oxygen, carbon and nitrogen. Furthermore, in the enriched gas cooling is further enhanced by common elements such as neon and iron.

With the exception of Compton cooling, all of the above processes require two particles;

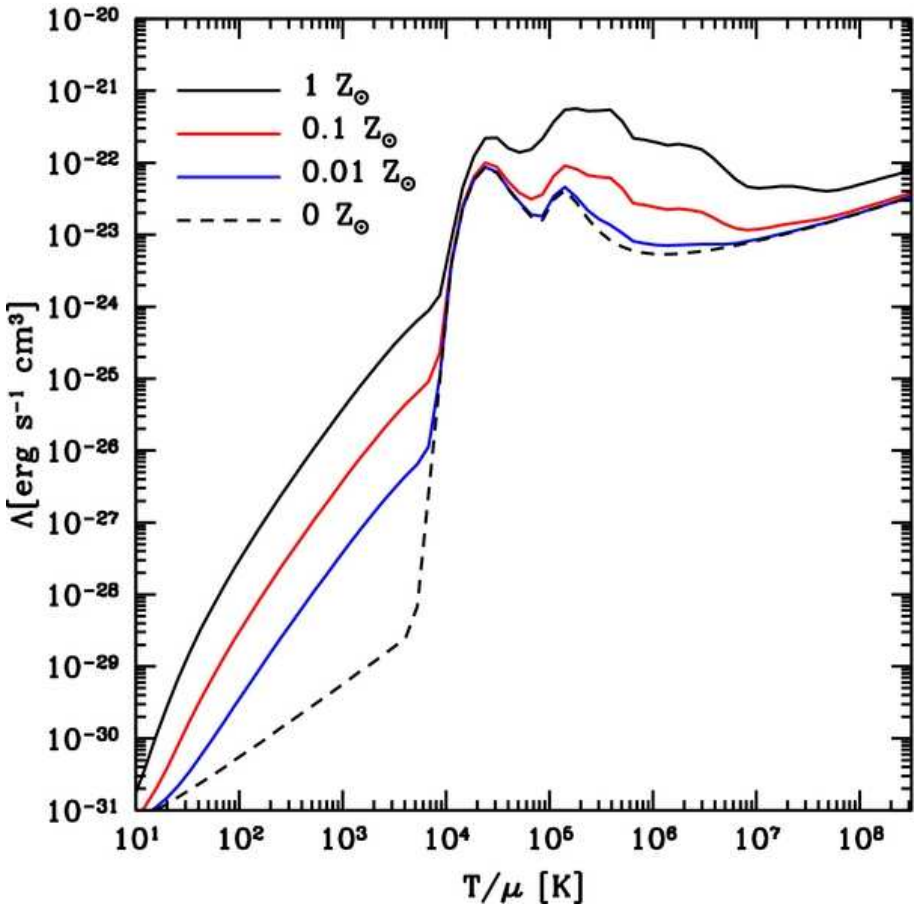

Figure 1.6: Effective cooling functions, in the absence of a UV background, implemented in the RAMSES code for varying metallicities, Z, as denoted in the legend in units of solar metallicity ($Z_\odot$). The peaks at temperatures $T \sim 15,000$ K and 10^5 K are due to the collisionally excited levels of H^0 and He^+ respectively, which are enhanced in the metal enriched gas due to fine line transitions of heavier elements such as oxygen, carbon and nitrogen. Other common heavy elements, such as neon and iron, are responsible for peaks in the cooling function at temperatures $T \gtrsim 10^5$ K, and the gradual rise in cooling efficiency at $T \gtrsim 10^7$ K is due to free-free emission (i.e.., Bremsstrahlung). *Credit:* Agertz et al. (2013).

therefore, their efficiency scales with gas density. At high densities, if the gas has been pre-enriched with metals, fine-line transitions become an effective cooling mechanism, dominating in the ISM. Photoheating plays a crucial role in these cooling processes. The Lyman-Werner bands photons (11.2 to 13.6 eV), produced in copious amounts by young massive stars, excite H_2 molecules, resulting in their dissociation and thus suppressing star formation in mini-haloes. As this process is thought to sterilize mini haloes (see e.g Ahn et al., 2009), we neglect detailed molecular chemistry in this thesis.

1.6.2 The First Stars and the Initial Mass Function

The vast majority of the observable Universe is due to emission from stars, however we do not have direct observational constrains on how the first stars formed at the end of the dark ages. In general, star formation is split into three episodic populations: the current day Population I stars, of which our Sun is an example, which form out of metal-rich gas and are mostly found in spiral galaxies, Population II stars which are relatively metal poor and typically found in globular clusters and the nuclei of galaxies, and finally Population III stars, which formed out of the primordial gas (H/He) and are extremely metal poor (Bromm and Larson, 2004). Due to their lack of metals, this first generation of stars is thought to have consisted of massive, extremely short lived stars with masses roughly in the region of $20 - 130\ M_\odot$ (Umeda and Nomoto, 2003). Their large masses correspond to large stellar luminosities, making them the prime candidates thought to have started the process of cosmological reionization.

While the physics of star formation is very complex, they are generally assumed to have formed out of cold, dense molecular clouds which for Population III stars have sizes comparable to the virial radius of their host halo, i.e $R_{\rm vir} \sim 100$ pc (Larson, 2003). For gas to cool in these haloes, their gravitational binding energy $\sim k_{\rm B}T_{\rm vir}$ per particle must exceed the gas thermal energy, $\sim k_{\rm B}T$ to retain accreted gas and prevent photoevaporation (Barkana and Loeb, 1999). At temperatures $T \lesssim 10^4$ K, the gas is limited to cooling via the rotational and vibrational lines of molecular hydrogen, for which the cooling time is smaller than the gas free fall time at $T \gtrsim 10^3$ K and molecular fraction $f_{\rm H_2} \sim 10^{-4}$ (e.g Haiman et al., 1996; Yoshida et al., 2003). Haloes with virial temperature $T_{\rm vir} \sim 10^3$ K therefore meet the criteria to allow cooling via molecular hydrogen, therefore are thought to host Population III stars. These so called *mini-haloes*, as their virial temperatures correspond to relatively low masses of $M_{\rm vir} \sim 10^6\ M_\odot$, are thought to have formed at $z \gtrsim 10$. As first confirmed by Machacek et al. (2001), a soft UV Ly-Werner background at

these times could delay the formation of these stars via the dissociation of H_2 molecules, thus denying haloes the only available cooling mechanism at these times, leaving a great deal of uncertainty regarding their formation and evolution.

One of the key questions today regards the form of the *primordial initial mass function* (IMF), which describes the distribution number of stars per logarithmic interval of stellar mass. The IMF of Population I stars is observed to take the approximate form of a Salpeter mass function:

$$\frac{\mathrm{d}N}{\mathrm{dlog}M_*} \propto M_*^{-x}, \tag{1.32}$$

where $x \simeq 1.35$ for $M_* \gtrsim 0.5\ M_\odot$ and $x \simeq 0.0$ for $M_* \lesssim 0.5\ M_\odot$ for a Kroupa (2002) IMF with a broken power law distribution. Other distributions also match observations well, such as the Chabrier (2003) log-normal distribution which differs only slightly from Kroupa (2002) at masses $\lesssim 1\ M_\odot$. The most important feature in the IMF of Population I stars is the characteristic mass scale which is $\sim 1\ M_\odot$, meaning that most of the mass goes into stars of similar masses to our sun. While it is assumed that the IMF is universal over time, there is no reason to believe that the IMF of the first stars should match that of present day Population I stars, given the drastic differences in their environment and chemical composition. The transition from Population III to II is thought to occur due to metal enrichment from supernovae (SNe) events. This is thought to have taken place prior to reionization ($z > 6$, Fang and Cen, 2004), an epoch for which observations are sparse. Therefore the IMF of Population III and II stars remains poorly determined.

1.6.3 Stellar Population Synthesis Models

Stellar population synthesis (SPS) models are tools for interpreting the integrated light observed from galaxies, based on a theoretical understanding of the properties of individual stellar populations. The starting point of any SPS model is the simple stellar population (SSP), which describes the evolution in time of the spectral energy distribution (SED) of a single population of stars (i.e the total energy emitted as a function of frequency) which form at the same time and according to a given IMF at fixed metallicity. The ingredients of any SPS therefore require a set of stellar evolutionary tracks (trajectories of stars in the Hertzsprung Russell diagram), stellar isochrones (loci in the Hertzsprung Russell diagram of stars with common age and metallicity) and finally a choice of IMF. A wide range of isochrone tables exist in the literature, which span a wide range of ages, chemical composition and evolutionary phases, for example the Padova (Bertelli et al., 1994; Girardi

et al., 2000; Marigo et al., 2008), BaSTI (Pietrinferni et al., 2004; Cordier et al., 2007), and Geneva models (Schaller et al., 1992; Meynet and Maeder, 2000). No single set spans the full range of ages, metallicities or evolutionary phases, therefore numerous SPS models are available in the literature, each with their own focus (e.g Bruzual A. and Charlot, 1993; Worthey, 1994; Devriendt et al., 1999; Fioc and Rocca-Volmerange, 1999; Leitherer et al., 1999; Girardi et al., 2000; Bruzual and Charlot, 2003). A commonality of these models is that they do not track the evolution of interacting binaries, hence stars are assumed to evolve independently of one another. In this scenario, hot, massive stars depart from the main sequence most rapidly, leading to a drop in UV flux with increasing age. However, as a significant fraction of stars are expected to belong to binary or multiple star systems, this could lead to uncertainties in SPS models.

1.7 Thesis Outline

The primary aim of this thesis is to understand the complex set of feedback processes which influence reionization. In particular, we probe the self-regulation of galaxies due to radiative feedback and the impact of the supersonic relative drift velocity between baryons and dark matter at recombination. To achieve this, we perform extremely high-resolution numerical simulations of galaxy formation in cosmological volumes, reaching scales in the tens-of-parsecs.

In Chapter 2, we outline the numerical methods and choice of simulation code including the calibration of sub-grid models for star formation and SNe feedback. We conclude this chapter by exploring potential alternatives to our star formation criteria, and briefly discuss the analysis pipeline developed to enable this work.

In Chapter 3, we investigate how the baryon fraction of dark matter haloes is modulated during reionization due to photoevaporation. This quantity is extremely important in semi-analytical models of galaxy formation, however is poorly understood. To better constrain this quantity, we develop and train an artificial neutral network to predict the baryon fraction and understand its dependence on reionization history.

In Chapter 4, we probe how stellar uncertainties surrounding the ionizing photon production rate of binary populations may impact the escape fraction of ionizing radiation. We discuss what implications this may have for the potential sources of reionization, and investigate the relationship between supernovae feedback and the escape fraction.

In Chapter 5, we present a new method for including the supersonic relative drift velocity between baryons and dark matter in volumes well above its coherence scale for

the first time. The drift velocity, a second-order term in the Euler equations of fluid dynamics, has until now only been included in very small volumes where these equations can be linearised. The advantage of much larger volumes lies in our ability to capture the statistical scatter between regions of varying relative velocity. We summarise the current state of this project, and discuss the plan of future work to follow.

Finally, we provide concluding remarks in Chapter 6 along with avenues for future work in 6.2.

Chapter 2

Simulating the Epoch of Reionization

2.1 Introduction

The promise of future observations combined with the relentless increase in computational power, has led to the introduction of radiative transfer in cosmological simulations. Relevant time-scales differ by orders of magnitude between the gas and radiation; the speed of light is much greater than the local sound speed, as such fully coupled radiative hydrodynamics (RHD) requires very short time steps to follow their evolution accurately. This is computationally expensive, with initial attempts limited to post-processing of radiative transfer to simulate reionization. While such implementations may successfully capture the global properties of reionization, such as its duration, the radiation is unable to influence the gas thus neglecting the impact on galaxy and star formation. Advancements in numerical algorithms and the dawn of peta-scale computing facilities have made fully coupled RHD possible, with the first state-of-the-art algorithms surfacing in recent years (Gnedin and Abel, 2001; Aubert and Teyssier, 2008; Wise and Abel, 2011; Rosdahl et al., 2013). The aim of this project is to simulate the EoR across a large dynamical range of scales, with a focus on the feedback on small-scales (galactic). To achieve this we require volumes large enough to capture a cosmologically representative fraction of the Universe, whilst simultaneously resolving the physics of sub-galactic scales such as star formation. Traditionally cosmological simulations solve the equations of fluid dynamics in either their Eulerian or Lagrangian forms. The advantages and drawbacks of these methods are well understood (for a review, see Agertz et al., 2007), however the latter is typically restricted to a fixed number of particles and therefore fixed spatial resolution of the gas. Eulerian

grid-based schemes typically use Adaptive Mesh Refinement (AMR) to alleviate this problem, as the initially coarse mesh is refined in regions of interest to achieve high spatial resolution at a greatly reduced cost. In this chapter, we discuss our choice of simulation code and contrast with possible alternatives.

2.2 Radiative Transfer Simulations

One method for simulating the EoR is using full cosmological simulations that attempt to incorporate all the relevant physics, including hydrodynamics, gravity and radiative transfer. While this method may be crucial in understanding the complex interplay of physical processes across a wide range of scales, particularly feedback, it poses significant computational challenges. The relevant scales during the EoR span from several Mpc to the scales of individual galaxies, and ultimately to the microscopic scale of individual atoms. While significant advances have been made in recent years, there exists no such miracle code which can self-consistently follow all of the above. Hydrodynamical effects typically limit the volume and spatial resolution of such simulations, as the numerical domain of dependence must contain the physical domain of dependence. For hydrodynamics, this arises from the finite time in which a pressure wave can propagate through the fluid element. This takes the form of a *Courant-Friedrichs-Levy* (CFL) condition, which can vary in time and limits the time-step, Δt, for a given spatial scale, Δx, such that the thermodynamical properties of the gas cannot change significantly in a single time-step. As a result, hydrodynamical simulations of reionization typically focus on small-scale effects, such as the self-feedback of galaxies and the escape fraction of ionizing radiation. A variety of radiative transfer algorithms are available in the literature, which appear to converge well for basic tests (e.g Iliev et al., 2006a, 2009). Radiative transfer poses many significant challenges, not least because it requires solving a seven-dimensional problem over time, position, frequency, and direction. For cosmological simulations, there is an added level of complexity as the number of sources increases exponentially with volume. Therefore cosmological radiative transfer codes typically take one of three approaches: (i) a Monte Carlo approach, in which a probabilistic method is used to trace large numbers of photon packets from each source (e.g Ciardi et al., 2001), (ii) ray tracing, in which rays are cast from each source to all elements of the computational domain (e.g Mellema et al., 2006a), or (iii) a field-based approach, where photons are treated as a continuous field and advected as a fluid (e.g Aubert and Teyssier, 2008). Each method has its own advantages and disadvantages, and scale differently depending on the precise nature of the

problem to be solved. For example, ray-tracing codes scale directly with the number of sources (and thus volume) while field-based methods scale independently of the number of sources, however are bound by a similar CFL condition to that of the hydrodynamics. Furthermore, they are quite diffusive which prevents the formation of shadows behind absorbers (Iliev et al., 2006a, 2009; Rosdahl et al., 2013).

2.3 Semi-numerical Simulations

The general agreement between the above methods has paved the way for hybrid semi-numerical schemes, which attempt to compromise between the accuracy of fully numerical techniques and the simplicity of analytical methods. The general procedure for any semi-numerical simulation is to first generate cosmological initial conditions, from which the density field can be evolved forward linearly by trivially scaling the initial amplitudes by the linear growing mode, $D_+(z)$, as introduced in section 1.3.1 (Loeb and Furlanetto, 2013). Potential sources are then catalogued, typically by identifying cosmological haloes, before finally generating the morphology of the ionized regions based on a sub-grid analytical approach. These approaches have the advantage that they can quickly and inexpensively explore the full parameter space of the underlying model for a detailed source distribution, which is often not possible with fully numerical methods. Semi-numerical techniques have been found to agree remarkably well with radiative transfer schemes, particularly on the predicted morphology of H II regions at the 10's of percent level. This suggests that these analytical models of reionization are quite robust (Zahn et al., 2011). Where these techniques potentially fail, however, is where the back-reaction of radiative feedback on gas is significant as the density and radiation fields are not evolved self-consistently. Another potential drawback is in the spherical modelling of H II regions, as hydrodynamical techniques have shown their growth to be quite anisotropic around ionizing sources (such as stars and QSOs), due to the geometry of dense filaments along which galaxies form (e.g Gnedin, 2000a; Shapiro et al., 2004; Feng et al., 2013). These methods are therefore best suited to probing the global statistics of reionization, whereas fully numerical techniques are well suited to probing galactic scales in small volumes.

2.4 Cosmological Initial Conditions

Cosmological initial conditions are generated by mapping fluctuations in the density fields of CDM and baryons, onto Gaussian random fields via their power spectra, introduced

in section 1.3.4 and following the initial methodology of Katz et al. (1994). In this approach, random Gaussian noise is sampled onto a uniform lattice in Fourier space, and then the amplitudes are scaled according to the input power spectrum, with velocity fields trivially constructed from the density field via the Fourier transform of the continuity equation. The power spectrum has the interesting property that it fully describes the statistical properties of any Gaussian random field, therefore provided there are no non-Gaussianities in the initial matter distribution, and fluctuations are still linear such that the linear growth factor remains valid, they can be used to accurately describe the initial conditions for structure formation. Fluctuations must however be large enough such that the integration error (shot noise) does not dominate over discretisation errors in constructing the continuous density fields. This is especially problematic on small-scales, where non-linearities first develop and are likely to be dominated by discreteness effects due to nearest-neighbour interactions (Baertschiger and Sylos Labini, 2002).

We generate initial conditions using the MUlti-Scale cosmological Initial Conditions (MUSIC) code (Hahn and Abel, 2013), which builds on the GRAFIC2 package for generating initial conditions with multiple levels of resolution, aptly named *zoom* initial conditions (Bertschinger, 2001). While the primary aim of MUSIC is to generate nested zoom regions, it can also be used to generate accurate uniform Cartesian initial conditions. In this thesis, we adopt both uniform and nested zoom initial conditions, where the latter are used to resolve scales where the supersonic baryon drift velocity modulates the matter power spectrum (see chapter 5).

2.5 Cosmological Radiation Hydrodynamics with RAMSES

2.5.1 Introduction

Obtaining a complete picture of the dark ages and the EoR requires that we simultaneously resolve scales of order a few hundred Mpc comoving and the Jeans mass. This corresponds to a dynamic range covering 9-10 orders of magnitude to capture reionization, a feat no code can achieve at this time. Nonetheless, numerical algorithms have undergone rapid advancements in recent years with the Illustris project recently achieving 6 orders of magnitude in dynamical range (Vogelsberger et al., 2014). For this work we have chosen the Eulerian tree-based N-body and hydrodynamics AMR code RAMSES (Teyssier, 2002), along with its RHD implementations ATON (Aubert and Teyssier, 2008) and RAMSES-RT (Rosdahl et al., 2013). Tree-based AMR schemes allow refinements on a cell-by-cell basis

(in contrast to the less efficient patch-based AMR), achieving a high spatial resolution in cosmological volumes. In the sections below, we briefly describe the set of equations which are solved in RAMSES for cosmological radiation hydrodynamics.

2.5.2 Dark Matter and the Poisson Equation

Dark matter particles are treated as a collisionless N-body system and governed by the Vlasov-Poisson equations

$$\frac{d\mathbf{x}_p}{dt} = \mathbf{v}_p \ , \quad \frac{d\mathbf{v}_p}{dt} = -\nabla\phi \ \text{ where } \ \nabla^2\phi = 4\pi G\bar{\rho} \tag{2.1}$$

where p denotes a given particle and ϕ the gravitational potential. To integrate the positions and velocities across time-steps, the density field is first computed from the discrete set of particles using a Cloud In Cell (CIC) interpolation scheme. The Poisson equation can then be solved to determine the gravitational acceleration on the mesh (Particle-Mesh, or PM method).

2.5.3 Hydrodynamics

The Euler equations are solved in their conservative form using a second-order Godunov method for a self-gravitating fluid,

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0 \tag{2.2}$$

$$\frac{\partial(\rho\mathbf{u})}{\partial t} + \nabla\cdot(\rho\mathbf{u}\otimes\mathbf{u}) + \nabla P = -\rho\nabla\phi \tag{2.3}$$

$$\frac{\partial E}{\partial t} + \nabla\cdot[(E+P)\mathbf{u}] = -\rho\mathbf{u}\cdot\nabla\phi + \Lambda(\rho,\varepsilon), \tag{2.4}$$

where:

ρ = gas density

t = time

$\mathbf{u}$ = bulk velocity of the fluid

P = pressure

ϕ = gravitational potential

E = total energy density (kinetic and thermal)

Λ = $\mathcal{H} - \mathcal{L}$ which denote radiative heating/cooling respectively.

The radiative heating/cooling term is separated in the final implementation, allowing for flexibility in the choice of thermochemistry solver. Generally this term is a function of the gas density, temperature, and ionization states of the gas. In RAMSES, collisional ionization equilibrium (CIE) is assumed for an optically thin plasma, such that the ionization states can be computed as functions of the gas density and temperature only (and are therefore not tracked quantities). Finally, the set of Euler equations is closed with an equation of state relating the pressure and internal energy:

$$E = \frac{1}{2}\rho u^2 + \varepsilon \tag{2.5}$$

$$P = (\gamma - 1)\varepsilon \tag{2.6}$$

which for a monatomic ideal gas, $\gamma = 5/3$. The gas temperature is related to internal energy by the ideal gas law:

$$\varepsilon = \frac{3}{2} n k_B T \tag{2.7}$$

where n is the number density of particles. Assuming a mean molecular mass $\mu m_{\rm H}$ we therefore have:

$$T = \frac{2}{3}\mu \frac{\varepsilon m_{\rm H}}{\rho k_B} \tag{2.8}$$

for hydrogen thermochemisty (H, H^+, e^-) only (addition of He in RAMSES-RT).

It is convenient to re-write the Euler equations as a generalized system of conservation laws:

$$\partial_t \mathcal{U} + \nabla \mathcal{F}(\mathcal{U}) = \mathcal{S} \tag{2.9}$$

where:

$\mathcal{U}$ $= [\rho, \rho\mathbf{u}, E]$ is the *conservative state vector*
$\mathcal{F}(\mathcal{U}) = [\rho\mathbf{u}, \rho\mathbf{u} \otimes \mathbf{u} + P, (E + P)\mathbf{u}]$ is the *flux vector*
$\mathcal{S}$ $= [0, -\rho\nabla\phi, -\rho\mathbf{u} \cdot \nabla\phi]$ is the *source vector*

neglecting the additional source term, Λ, as this is separated in the numerical implementation. The vector $\mathcal{U}$ denotes the cell averaged quantities tracked in RAMSES, with the addition of a ρZ term when tracking metallicity as a simple scalar quantity. For a detailed explanation of the discretization of equation (2.9), see Teyssier (2002).

2.5.4 Radiative Transfer

Coupled radiative transfer (RT) in RAMSES can be performed using either the ATON or RAMSES-RT modules. Both use the same moment based radiative transfer scheme with the M1 closure to the Eddington tensor, as outlined in Aubert and Teyssier (2008). This scheme reduces the equation of radiative transfer to a set of hyperbolic conservation laws, which are analogous to treating the radiation field as a fluid with a bulk flow. In contrast, traditional ray-tracing codes cast rays from each source through each cell, which scales with the total number of ionizing sources. The M1 moment method therefore has the advantage that it scales completely independently of the number of sources, as each source deposits its ionizing photons in the nearest grid cell from which it is advected to all neighbouring cells. The ATON module uses graphical processing units (GPUs) to perform RT as a post-processing step after each hydro time-step, and is coupled to the hydrodynamics with non-equilibrium thermochemistry of hydrogen gas only.

To achieve the speed-up to perform RT at relatively low cost, ATON requires well-structured Cartesian grids, so that cells are arranged optimally in both space and memory. As a result, ATON is not well suited to AMR, instead the coarse level grid is projected onto a Cartesian grid to evolve the transfer of photons and disabling AMR in the process. For cosmological volumes this restricts us somewhat in resolution, therefore ATON is well suited to large-scale reionization problems opposed to small-scale galaxy formation. On the contrary, RAMSES-RT solves the radiative transport step on the CPU and is coupled via hydrogen and helium theremochemistry. The code is therefore fully integrated into the existing hydrodynamics solver, achieving fully-coupled and self-consistent radiation hydrodynamics. The cost of this implementation, is that the radiative transfer step becomes expensive due to the number of sub-cycling steps required, for which the CPU is not as efficient as the GPU. However, we regain the AMR capability allowing for much higher resolutions. We briefly outline the M1 moment scheme for radiative transfer here for completeness, following the description by Aubert and Teyssier (2008); Rosdahl et al. (2013).

To begin, we define the radiative transfer equation (Mihalas and Mihalas, 1984):

$$\frac{1}{c}\frac{\partial I_\nu}{\partial t} + \mathbf{n}\cdot\nabla I_\nu = -\kappa_\nu I_\nu + \eta_\nu, \tag{2.10}$$

where $I_\nu(\mathbf{x},\mathbf{n},t)$ is the radiation specific intensity, $\kappa_\nu(\mathbf{x},\mathbf{n},t)$ is the absorption coefficient, $\eta_\nu(\mathbf{x},\mathbf{n},t)$ is the source function, $\mathbf{n}$ is the direction of propagation, and c is the speed of light. This equation is analogous to the continuity equation of fluid dynamics, and de-

scribes the conservation of photons in phase space. The absorption coefficient is simply the product of the neutral hydrogen number density and the cross-section to photoionization, i.e $\kappa_\nu = n_{\rm HI}\sigma_\nu$. Assuming isotropic sources, we can average over solid angle to define the *angular moments* of the specific intensity. The zeroth moment is trivial, and defines the mean intensity of differential energy density:

$$4\pi J_\nu = \oint I_\nu \, \mathrm{d}\mathbf{\Omega} = cE_\nu(\mathbf{x}, t), \tag{2.11}$$

where $E_\nu(\mathbf{x}, t)$ is the radiation energy, and $\mathrm{d}\Omega$ denotes an integral over solid angle. The first and second moments correspond to the radiation flux and momentum pressure tensor:

$$\mathbf{F}_\nu(\mathbf{x}, t) = \oint I_\nu(\mathbf{x}, \mathbf{n}, t)\mathbf{n} \, \mathrm{d}\mathbf{\Omega}, \tag{2.12}$$

$$\mathbb{P}(\mathbf{x}, t) = \frac{1}{c} \oint I_\nu(\mathbf{x}, \mathbf{n}, t)\mathbf{n} \cdot \mathbf{n} \, \mathrm{d}\mathbf{\Omega}. \tag{2.13}$$

Using equation 2.10, these moments satisfy a hyperbolic set of conservation laws for the radiation fluid:

$$\frac{\partial E_\nu}{\partial t} + \nabla \mathbf{F}_\nu = -\kappa_\nu c E_\nu + S_\nu, \tag{2.14}$$

$$\frac{\partial \mathbf{F}_\nu}{\partial t} + c^2 \nabla \mathbb{P}_\nu = -\kappa_\nu c \mathbf{F}_\nu, \tag{2.15}$$

where S_ν is the source term and drops out in the second moment due to isotropy. Finally, by dividing equations 2.14 and 2.15 by a single photon energy, $h_p\nu$, the energy and flux densities can by replaced by number densities:

$$\frac{\partial N_\nu}{\partial t} + \nabla \mathbf{F}_\nu = -\kappa_\nu c N_\nu + S_\nu, \tag{2.16}$$

$$\frac{\partial \mathbf{F}_\nu}{\partial t} + c^2 \nabla \mathbb{P}_\nu = -\kappa_\nu c \mathbf{F}_\nu, \tag{2.17}$$

where N_ν is the photon number density, and for the sake of simplicity we have used the same notation for $\mathbf{F}_\nu$ and $\mathbb{P}_\nu$ despite differing by a factor $h_{\rm p}\nu$. This system of equations resembles the Euler equations for fluid dynamics, hence represent a set of hyperbolic conservation laws for the photon fluid.

Moment-based schemes for RHD are advantageous over ray-tracing codes for several reasons; by adopting a fluid treatment, radiation is not tracked from individual sources.

Rather, photons are injected into the fluid and evolved across cell boundaries. This does however, have an obvious drawback that multi-directionality is lost, and replaced with a bulk flow of the fluid leading to a somewhat diffusive field. For isotropic sources (such as stars) this is not an issue, and therefore for optically thick discs (galaxies) this is a good description. Furthermore, the radiation field in each cell can be time-integrated using only local quantities. In both RAMSES-RT and ATON, the size of the RT time-step is inversely proportional to the speed of light, as constrained by the Courant condition.

To reduce load, one can adopt a reduced speed of light approximation, as proposed by Gnedin and Abel (2001), to relax the Courant condition. This approximation is discussed in detail in RAMSES-RT, and a framework is offered for determining the appropriate value based on relevant time-scales. For galaxy formation problems, a reduced speed of light can still arrive at the correct solution, however large-scale reionization problems are more complicated. This code achieves convergence for the fourth test of Iliev et al. (2006a), ionizing a cosmological volume, only if the full speed of light is used. Therefore the code is not well suited to studying the global evolution of reionization, but rather the internal properties of haloes and galaxies.

Relevant particle interactions such as photoionization, collisional (recombination, ionization, excitation cooling), Bremsstrahlung cooling and Compton heating/cooling are included in both codes and therefore we no longer assume CIE. The source term is split and solved separately in an injection step, followed by a transport step and finally the thermochemical step. For the transport step, the right hand side of equations 2.16 and 2.17 are set to zero, and the conservative and flux state vectors updated for free-flowing photons:

$$\mathcal{U}_{\rm rad} = [N_\nu, \mathbf{F}_\nu], \tag{2.18}$$

$$\mathcal{F}_{\rm rad}(\mathcal{U}) = [\mathbf{F}_\nu, c^2\mathbb{P}_\nu], \tag{2.19}$$

where for RAMSES-RT the equations are discretized over frequency into m photon groups which are solved independently for hydrogen and helium thermochemistry. For cosmological simulations, "super-comoving" coordinates are adopted (Martel and Shapiro, 1998) which involves a change of variables for each element in $\mathcal{U}$ to its comoving analogue. Hence, equation 2.9 can be used in its current form for both standard fluid-dynamics and cosmological simulations (for details on their numerical discretization, see Aubert and Teyssier, 2008; Rosdahl et al., 2013). We used the ATON code coupled with RAMSES to

complete the largest fully-coupled radiation hydrodynamics simulation to date in Ocvirk et al. (2016), while RAMSES-RT is used to test the findings of this work on galactic scales, using extremely high dynamical range AMR simulations in this thesis.

2.6 Dynamical Range

The primary advantage of AMR codes over Smoothed Particle Hydrodynamics (SPH), such as the GADGET-2 code (Springel, 2005), is their ability to achieve high dynamical ranges by refining the Eulerian grid. Similar methods exist for SPH based codes, which involve costly particle splitting and merging on-the-fly (e.g Kitsionas and Whitworth, 2002; Vacondio et al., 2013). Several popular AMR cosmolgical codes currently exist, such as ART (Kravtsov et al., 1997), ENZO (Bryan et al., 2014), and the above discussed RAMSES code (Teyssier, 2002). The principles are the same in each, however the implementation can vary. For example, ENZO uses a patch-based refinement scheme, where the smallest grids which can be refined consist of a rectangular grid of cells, while RAMSES supports tree-based refinement, where each cell can individually be refined as per user defined criteria, based on the fully threaded tree algorithm of Khokhlov (1998). The latter is better suited to the hierarchical clustering problem posed by simulations of galaxy formation, where a highly complex geometry can build up and the number of grids which would be required by patch-based codes would be far less efficient.

For this reason we have chosen to use RAMSES for this thesis as discussed in previous sections. To prove our capability to reach extremely high resolution and dynamical range, we show an example projection of an individual halo taken out of a simulated cosmological density field in figure 2.1, taken from our RT2 simulation presented in chapter 3. The boxsize is set to 4 Mpc h^{-1} comoving, and the effective resolution is $\Delta x = 21.72$ pc, corresponding to a dynamical range of $\sim 2 \times 10^5$ which is close to that achieved by Vogelsberger et al. (2014).

2.7 Star Formation and Feedback: Calibration

2.7.1 Star Formation

The highest resolution simulations of galaxy formation in cosmological contexts are still limited to parsec scales, far larger than the scale at which the formation of a single star takes place. Therefore, a sub-grid recipe is implemented in RAMSES, which follows a

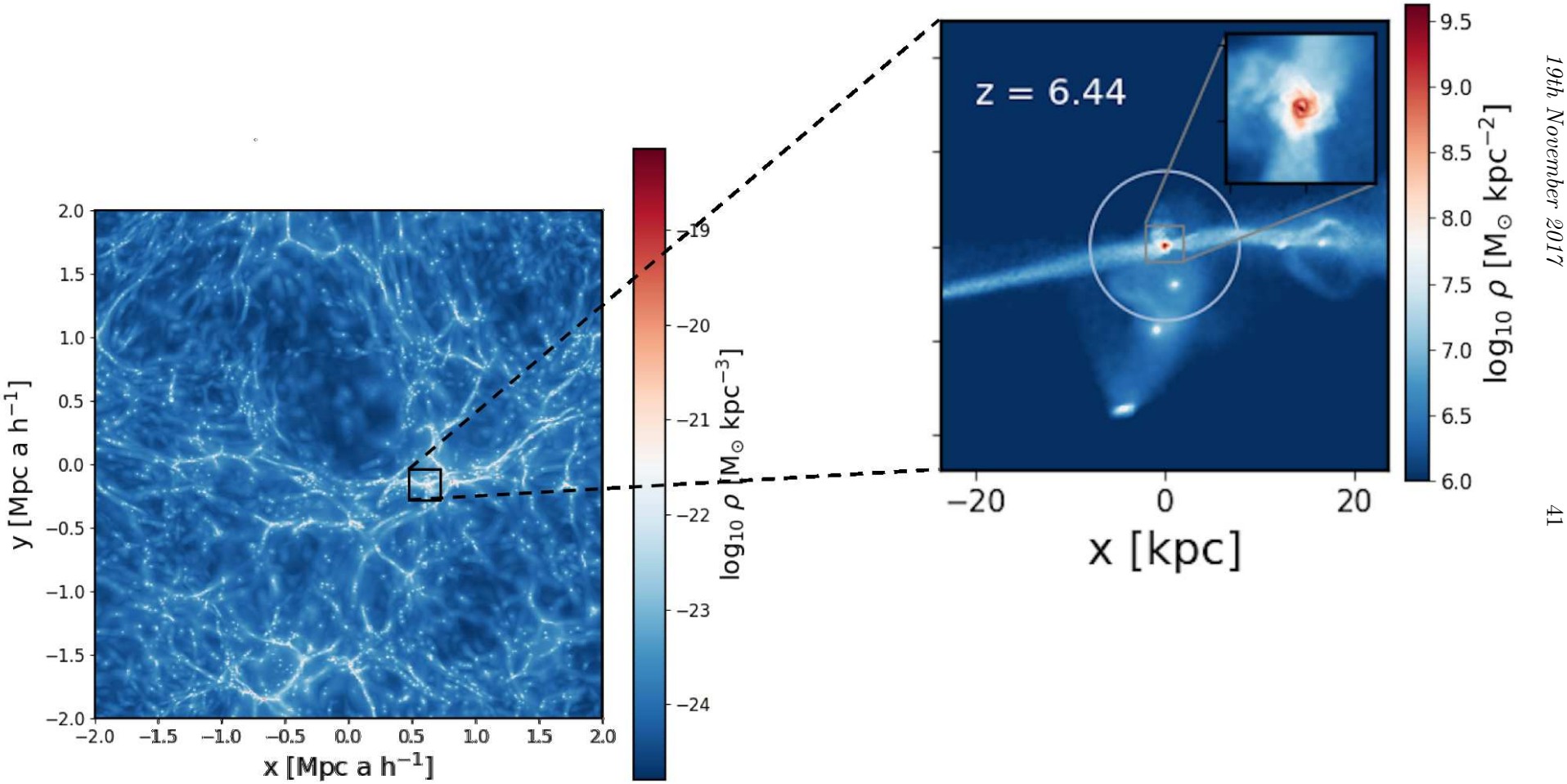

Figure 2.1: Slice of a cosmological density field at $z = 6.44$ (left) and a projection of a single halo (right) to demonstrate the dynamical range in our RT2 simulation presented later in chapter 3. The light blue circle in the right hand panel denotes the virial radius of the halo, and we zoom further into the central galaxy to demonstrate our ability to resolve spiral features for an isolated disc galaxy in a cosmological volume.

standard Schmidt law in each cell:

$$\dot{\rho}_* = \epsilon_{\rm ff}\rho/t_{\rm ff}, \tag{2.20}$$

where $\epsilon_{\rm ff} = 0.01$ is the star formation efficiency per free fall time $t_{\rm ff} = [3\pi/(32G\rho)]^{1/2}$, and G is the gravitational constant. Cells are considered star-forming when they meet the predefined density and temperature criteria:

$$\begin{aligned} n_* &\geq 8 \text{ cm}^{-3} \\ T/\mu &\leq 2 \times 10^4 \text{ K}, \end{aligned} \tag{2.21}$$

where n_* is the density threshold for star formation in units of the number of hydrogen atoms per cubic centimetre, and μ is the mean molecular weight of the gas. When a cell becomes eligible, N_* collisionless star particles are formed stochastically, with probability drawn from a Poisson distribution:

$$P(N_*) = \frac{\lambda_{\rm P}}{N_*!}\exp(-\lambda_{\rm P}), \tag{2.22}$$

with mean value

$$\lambda_{\rm P} = \left(\frac{\rho\Delta x^3}{m_*}\right)\frac{\Delta t}{t_{\rm ff}}, \tag{2.23}$$

where Δx is the minimum cell width and Δt is the simulation time step (Rasera and Teyssier, 2006). The stellar particle mass is then determined as an integer multiple of $m_* = m_{\rm H} n_* \Delta x^3/(1 + \eta_{\rm SNe} + \eta_{\rm w})$, where $m_{\rm H}$ is the mass of a hydrogen atom, and the parameters $\eta_{\rm SNe}$ and $\eta_{\rm w}$ are discussed in Section 2.7.3.

2.7.2 Calibration

An artificial 'Jeans pressure' is imposed on the gas to prevent numerical fragmentation below the Jeans scale of fine grid cells (Truelove et al., 1997), which can be written as:

$$\lambda_{\rm J} = \sqrt{\frac{\pi c_{\rm s}}{G\rho}} = 16 \text{ pc}\left(\frac{T}{1 \text{ K}}\right)^{\frac{1}{2}}\left(\frac{n_{\rm H}}{1 \text{ cm}^{-3}}\right)^{-\frac{1}{2}}, \tag{2.24}$$

where $c_{\rm s} = \sqrt{\gamma k_{\rm B} T/m_{\rm p}}$ is the sound speed, and we assume a ratio of specific heats $\gamma = 5/3$ for a monoatomic gas. Imposing the condition that the Jeans length be resolved on N cells gives rise to a temperature floor, which must satisfy (Rosdahl et al., 2015):

$$\frac{T}{1\text{K}} \geq \frac{n_{\rm H}}{1 \text{ cm}^{-3}}\left(\frac{N\Delta x}{16 \text{ pc}}\right)^2. \tag{2.25}$$

We apply this in the form of an effective temperature function (added to the cell temperature) with polytropic equation of state:

$$T_{\rm J} = T_0\left(\frac{n_{\rm H}}{n_*}\right)^{g_*-1}, \tag{2.26}$$

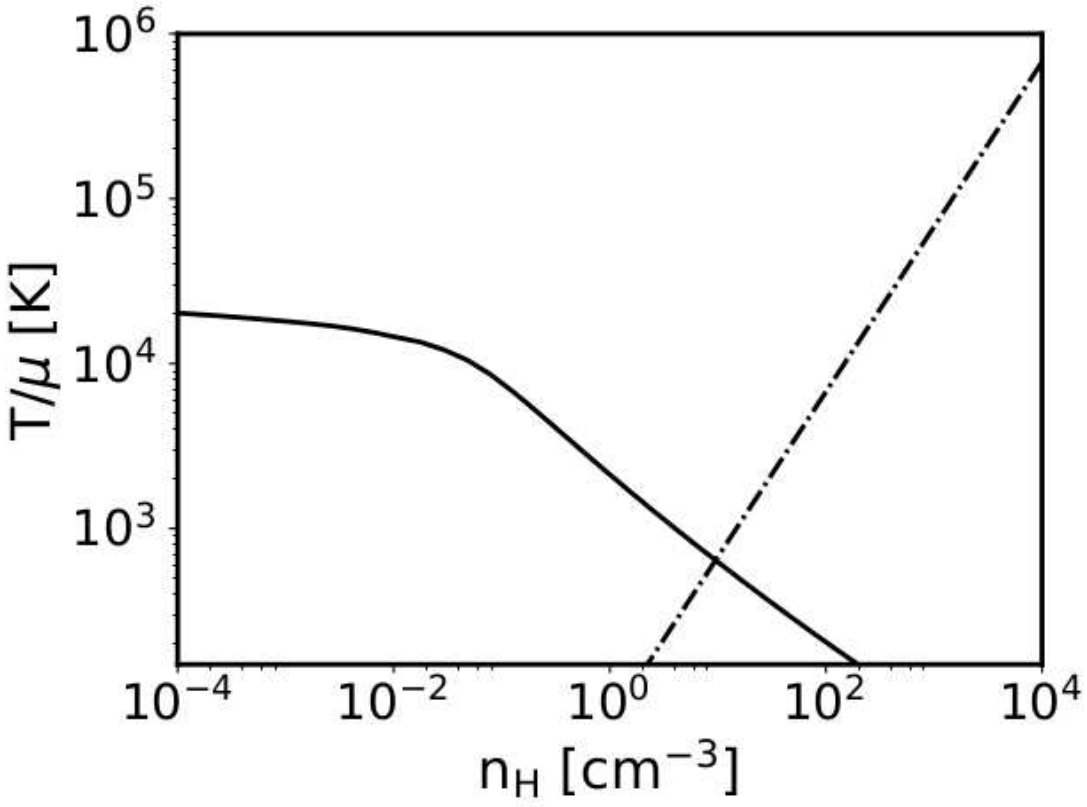

Figure 2.2: Temperature-density relation (solid) and intersection with our polytropic temperature floor (dot-dashed) used to calibrate star formation criteria. The solid line shows the trajectory in the temperature-density plane for a single cell at solar metallicity, as calculated using the RAMSES cooling module.

where g_* is the polytropic index (set to 2 in our models). Note that as this pressure floor is non-thermal, the gas temperature T can fall below this floor, therefore, the quantity $T - T_\mathrm{J}$ must satisfy the temperature criterion in equation (5.25). In Fig. 2.2, we show our polytropic equation of state and its intersection with the equilibrium cooling curve, which we calculate using the cooling module from RAMSES. The cooling module implements thermal Bremsstrahlung, ionization, recombination, dielectric recombination, and metalline cooling, as-well as radiative and Compton heating. We compute the temperature-density relation for a single cell at solar metallicity, and proceed to calibrate by adjusting n_* and T_0 such that the intersect point remains fixed for a given resolution.

2.7.3 Supernovae Feedback

We incorporate SNe using the kinetic feedback implementation of Dubois and Teyssier (2008), which models the Type II SNe explosion by increasing the kinetic energy of the

surrounding gas within a spherical blast wave of radius $r_{\rm SNe}$. Once a cell becomes star-forming, gas mass is depleted and converted into star particles such that

$$(\Delta m_{\rm g})_{\rm SF} = m_*(1 + \eta_{\rm SNe} + \eta_{\rm w}), \tag{2.27}$$

where m_* is the star particle mass, $\eta_{\rm SNe}$ is the SNe mass fraction, and $\eta_{\rm w}$ is the mass-loading factor (i.e. the amount of gas entrained within the SNe ejecta; Springel and Hernquist 2003). A value of $\eta_{\rm SNe} = 0.2$ is chosen, which roughly corresponds to a Chabrier (2003) stellar IMF. Observational evidence suggests the mass-loading factor varies with stellar mass (Martin, 2005; Rupke et al., 2005; Weiner et al., 2009; Martin et al., 2012; Newman et al., 2012; Chisholm et al., 2015; Arribas et al., 2014); however, a value $\eta_{w,\rm max} = 10$ is acceptable and therefore used here. The SNe explosion is treated as a single event 10 Myr after the formation of the star particle, releasing kinetic energy:

$$E_{\rm d} = \eta_{\rm SNe}\frac{m_*}{M_{\rm SNe}}E_{\rm SNe} \tag{2.28}$$

into the neighbouring gas cells, where $M_{\rm SNe}$ and $E_{\rm SNe}$ are fixed for a typical Type II SNe explosion (10 $\rm M_\odot$ and 10^{51} erg, respectively). In practice, the final mass-loading factor can be as low as zero, depending on the density of gas in neighbouring cells at the point of explosion. The most important parameter in the model is the radius over which the fluid variables are modified, $r_{\rm SNe}$, which we choose to be 200 pc and is consistent with the predicted size of galactic super-bubbles (McKee and Ostriker, 1977). While this scheme is not as sensitive to the over-cooling problem as thermal dump methods, it is also not immune: kinetic energy is converted into heat via shocks immediately after the SNe explosion. Given the typically high Mach number, ambient density, and limited numerical resolution, this heat can be efficiently radiated away, reducing the overall impact of SNe feedback (Kimm and Cen, 2014). As SNe are not thought to play an important role in the stripping of baryons from haloes, we do not expect this to significantly effect our results (Dubois and Teyssier, 2008; Roškar et al., 2014).

2.7.4 Stellar Population Synthesis and UV Background

As the emission of ionizing photons from stellar sources is central to much of the work in this thesis, the choice of SPS has a major impact on our results. Where not otherwise stated, we use the SPS model of Bruzual and Charlot (2003) (hereafter BC03) when interpolating stellar luminosities, assuming a Chabrier (2003) IMF, as presented in the left hand panel of figure 2.3. The vertical lines correspond to the ionization wavelengths of HI, HeI, and HeII, which mark the mean wavelengths of each photon bin used, as RAMSES-RT

supports multi-frequency RT. While the BC03 models span a large range of frequencies at high resolution, they neglect a detailed physical treatment of binary populations. This will become important in chapter 4, where we will discuss the implications and alternatives in greater detail. Here we show the evolution in stellar age of a single SSP at solar metallicity, $Z_\odot$. The full range of tabulated ages spans the range 0 Myr$-$2 Gyr, across 221 bins evenly space logarithmic bins, each with metallicities $1 \times 10^{-4} Z_\odot - 5 \times 10^{-2} Z_\odot$ over 6 evenly spaced logarithmic bins. Finally, wavelength is tabulated into 6900 bins spanning the range $91\text{Å} \leq \lambda \leq 1.6 \times 10^6\text{Å}$. The majority of ionizing photons are emitted in the first ~ 3 Myr of stellar evolution, with the UV luminosity decreasing with age as massive UV bright stars begin to die off.

Age-metallicity dependent population luminosity, L, is then computed as described in Rosdahl et al. (2013), by:

$$L_i(\tau_{\rm SSP}, Z) = \int_{\nu_{i,0}}^{\nu_{i,1}} J_\nu(\tau_{\rm SSP}, Z)/h_p \nu d\nu, \tag{2.29}$$

where h_p denotes Plancks constant, $\tau_{\rm SSP}$ and Z are the population age and metallicity, and $J_\nu = c/\nu^2 J_{\lambda(\nu)}$ with $J_{\lambda(\nu)}$ being the quantity shown in the right hand panel of figure 2.3, and $\nu_{i,0}$ and $\nu_{i,1}$ denote the lower and upper limits of the frequency bin belonging to photon group i. The cumulative luminosity is then given by the integral over time:

$$\Pi_i(\tau_{\rm SSP}, Z) = \int_0^{\tau_{\rm SSP}} L_i(t, Z) dt. \tag{2.30}$$

These quantities are tabulated at the beginning of each simulation run, to prevent them having to be recomputed at each photon injection step.

Prior to fully-coupled radiation hydrodynamical simulations of reionization, several authors (e.g Gnedin, 2000b; Hoeft et al., 2006; Okamoto et al., 2008) resorted to modelling UV emission from stellar sources as a (quasi)-homogeneous soft UV background, where reionization is assumed to occur instantaneously at some pre-determined redshift. Such a model can also be included in RAMSES-RT, to account for background spectrum of star-forming galaxies and quasars, assuming the latter contributes a significant number of photons to the HI ionizing field. We include this background spectrum in the form of a quasi-homogeneous, redshift evolving UV background using the model of Faucher-Giguère et al. (2009). When used, we initialise the background radiation field at $z = 9$ and evolve it in time as shown in the right hand panel of figure 2.3. Under this approximation, all cells become emitters, and photons are released into their corresponding photon groups as discussed above. The intensity of the UV background field initially grows with the

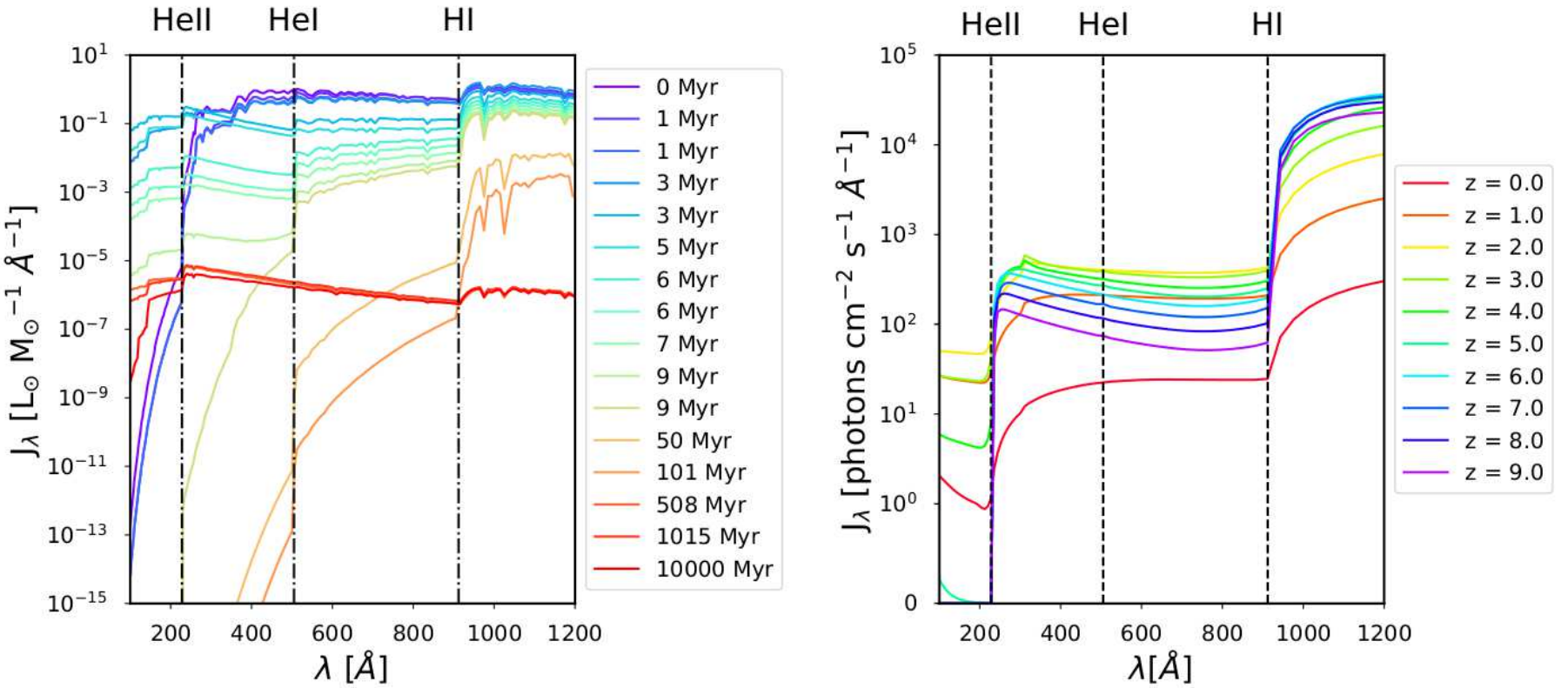

Figure 2.3: *Left panel:* Spectral Energy Distribution (SED) for a SSP at solar metallicity of Bruzual and Charlot (2003), across the age range 0 Myr − 10000 Myr and solar metallicity, $Z_\odot$. Vertical lines denote the ionization wavelengths of HI, HeI, and HeII which mark the mean wavelengths of each photon group in our simulations. The shape and intensity of the spectrum evolves in time due to the post main-sequence evolution of stars. *Right panel:* Redshift evolution of the homogeneous UV background model of Faucher-Giguère et al. (2009), used to model the background spectrum of star-forming galaxies and quasars. The field is initialised at a pre-determined redshift, $z = 9$ in our case, and all cells are thereafter assumed to be optically thin to ionizing UV and hence become emitters. The intensity initially increases due to the exponential growth of the source population, before turning over at $z \sim 2$ due to the accelerated expansion of the Universe. Where not otherwise stated, these models are used as the defaults in this work.

exponential growth of the source population, before turning around at $z \sim 2$ and beginning to decline due to the accelerated expansion of the Universe.

2.8 Alternative Star Formation Models

As discussed above, the sub-grid star formation model available in RAMSES is heavily resolution dependent, and requires careful calibration. The exact combination of parameters which results in a good calibration is also highly arbitrary, which coupled with the computational expense of coupled RHD simulations (even relatively small tests) makes it impractical to properly explore the available parameter space. In this section, we present an alternative sub-grid star formation model, based on the self-gravitating and resolution independent criterion of Hopkins et al. (2013). This model has no resolution dependent tunable parameters, potentially simplifying the calibration process significantly. We briefly describe the model below, and present preliminary results for the star formation rate in an isolated disc simulation.

On some scale δ_r, we require $\sigma^2_{\rm eff} + c_s^2 < \beta GM(< \delta_r)/\delta_r$, where $\sigma_{\rm eff}$ is the effective velocity dispersion which includes the contributions from both rotational and random motions, and β is a constant which depends on the internal structure. In practice such gas is always highly supersonic in simulations (and observations), so the c_s term can be neglected. We then arrive at the usual virial parameter:

$$\alpha \equiv \sigma^2_{\rm eff}\delta_r/\beta GM(< \delta_r) \tag{2.31}$$

Taking the limit where δ_r is small, we can write $M(< \delta_r) = (4\pi/3)\bar{\rho}\delta_r^3$ where $\bar{\rho}$ is the mean density in δ_r, and $\sigma_{\rm eff} \longrightarrow \frac{\delta_{\rm v}}{\delta_{\rm r}}\delta_r$, or

$$\sigma^2_{\rm eff} = \beta_v(|\bigtriangledown \cdot \mathbf{v}|^2 + |\bigtriangledown \times \mathbf{v}|^2)\delta_r^2 \equiv \left(\frac{\delta_v}{\delta_r}\right)^2 \delta_r^2 \tag{2.32}$$

The $|\bigtriangledown \cdot \mathbf{v}|$ term accounts for the local radial velocity dispersion and inflow plus outflow, while the $|\bigtriangledown \times \mathbf{v}|$ term accounts for internal rotational, shear, and tangential dispersion. The constant β_v again depends on the internal structure but is close to unity. Combining these terms we arrive at the resolution independent criterion:

$$\alpha \equiv \frac{\beta'}{2}\frac{|\bigtriangledown \cdot \mathbf{v}|^2 + |\bigtriangledown \times \mathbf{v}|^2}{G\rho} < 1, \tag{2.33}$$

where we assume the gas collapses in a single free-fall time ($\epsilon_{\rm ff} = 1$ for a homogeneous sphere, approximating individual cloud results of Padoan et al., 2012).

We test this model by initialising an isolated disc with two idealised free-falling streams feeding star forming gas, using the DICE code (Perret, 2016), and shown in figure 2.4 where

the colour-bar denotes the projected gas density along the line of sight. Here, we have rotated the disc to be side-on to show the streams. The total boxsize is set to 300 kpc and the circular velocity of the host halo is $V_c = 200$ km s^{-1}, corresponding to a virial mass of $M_{\rm vir} = 1 \times 10^{10}\ M_{\odot}$. We let the AMR hierarchy refine up to level $\ell_{\rm max} = 12$, corresponding to a formal resolution of $\Delta x = 73.24$ pc. The refinement criteria used is based on a quasi-Lagrangian strategy, similar to that of our cosmological simulations presented later in this thesis. We allow cooling via metal fine-line transitions down to $T \sim 100$ K in both models. The disc is initially devoid of stars, from which we allow stars to form based on the default density and temperature criteria of RAMSES (see section 2.7.1) and the above criterion of Hopkins et al. (2013). We fix the stellar mass in both cases, such that it roughly corresponds to 1/10 the dark matter particle mass for this volume and resolution. Finally, we allow SNe feedback as described in section 2.7.3.

The resulting star formation rates (SFR) for both models are plotted in figure 2.5, which show reasonably good agreement for the first 200 Myr of evolution. The Hopkins et al. (2013) initially lags behind the default RAMSES case, and is also more sensitive to SNe feedback as demonstrated by the bursty SFR pattern. Supernovae feedback eventually begins to regulate the total SFR, as seen after $t \sim 175$ Myr. While the Hopkins et al. (2013) method is in principle a more physical and robust criterion, computing the derivatives of the velocity field for each cell in the simulation domain poses a significant bottleneck on performance. This can be somewhat alleviated by imposing a lower limit on the density of star forming cells, however it is still impractical for cosmological volumes. To compute the derivatives accurately for each galaxy in such a volume would require a refinement criterion which is optimised to resolve the ISM. One possibility could be to ensure that the Jeans length of gas above typical ISM densities, $n_{\rm H} \sim 1{\rm cm}^{-3}$, is always resolved on a fixed number of cells (typically the same number used in our calibration; see section 2.7.2). Again, this is quite computationally expensive, especially for larger volumes, however could in principle be used in cosmological simulations. While this result allows us to conclude that our method of calibration provides a good normalisation, however, additional testing beyond the scope of this thesis is required to truly determine the accuracy of both models.

2.9 Analysis Pipeline

For the purpose of this thesis, a new analysis pipeline was developed designed to work with large AMR RAMSES datasets, called SEREN3. The package was designed from the ground up to take advantage of distributed memory computing clusters, supporting both OpenMP

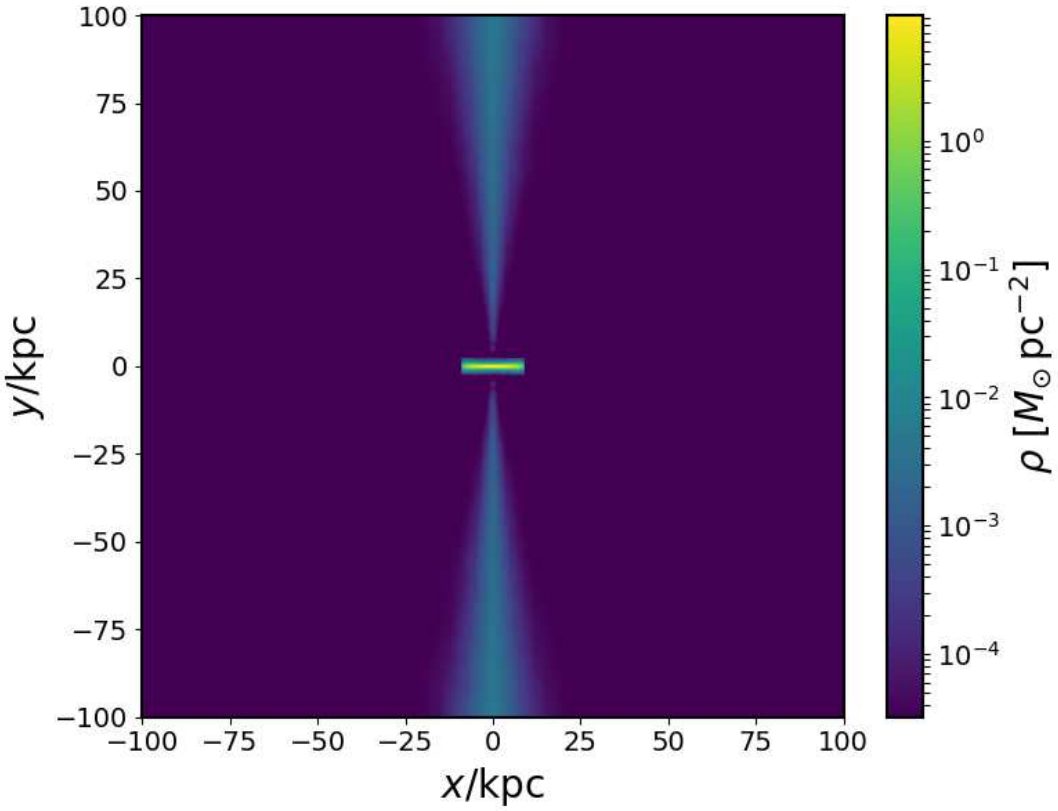

Figure 2.4: Side-on projection of our isolated disc initial conditions with two idealised streams, generated with the DICE code (Perret, 2016). The total boxsize is set to 300 kpc and the AMR hierarchy is allowed to refine up to level 12, leading to a formal resolution $\Delta x = 73.24$ pc. We allow stars to form using two methods: the standard density and temperature criterion in RAMSES, and the self-gravitating condition of Hopkins et al. (2013).

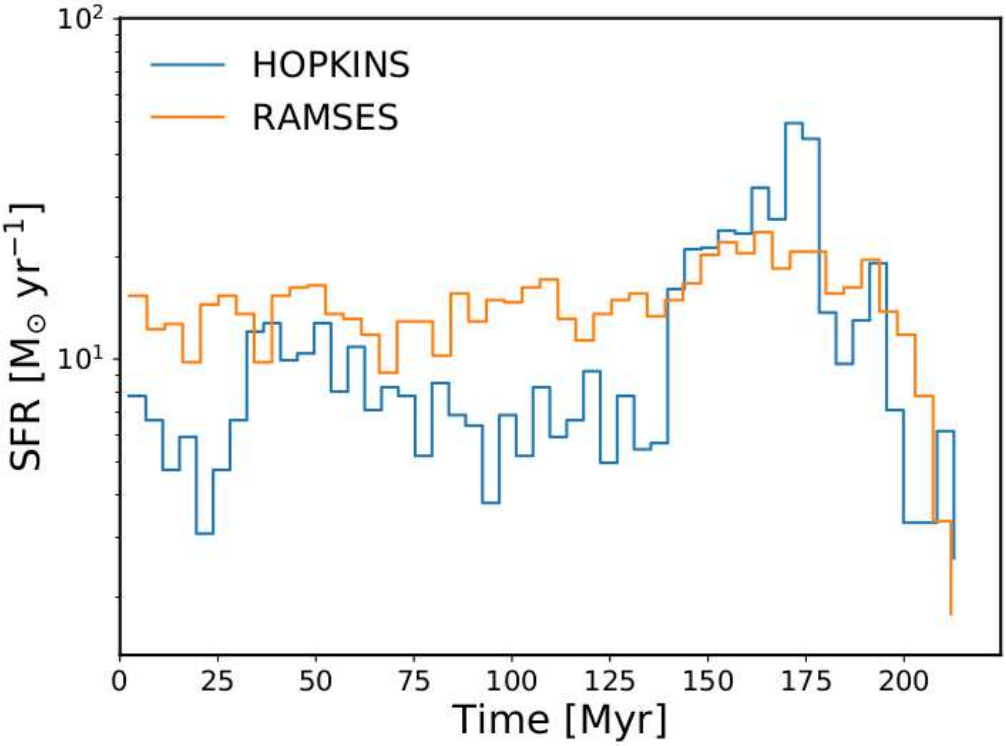

Figure 2.5: Side-on projection of our isolated disc initial conditions with two idealised streams, generated with the DICE code (Perret, 2016). The total boxsize is set to 300 kpc and the AMR hierarchy is allowed to refine up to level 12, leading to a formal resolution $\Delta x = 73.24$ pc. We allow stars to form using two methods: the standard density and temperature criterion in RAMSES, and the self-gravitating condition of Hopkins et al. (2013).

and MPI parallel processing, and is highly object orientated. At its core, data I/O of quantities tracked by RAMSES (i.e quantities tracked in the conservative state and flux vectors, $\mathcal{U}$ and $\mathcal{F}(\mathcal{U})$, in section 2.5) is handled by the excellent PYMSES module (Guillet et al., 2013), while SEREN3 implements the dynamically derived fields, unit tracking and conversion, and halo catalogue loading and filtering.

Deriving new fields in SEREN3 is a trivial one-time process. All that is required is for the user to define a new method with the correct method signature and annotation, as shown by the example below for gas mass, and the code caches the function and all relevant information on initial import to derive the field whenever needed.

```
import seren3

@seren3.derived_quantity(requires=["rho", "dx"])
def amr_mass(context, dset, **kwargs):
    '''
    Return cell mass
    '''
    rho = dset["rho"]  # kg m^-3
    dx = dset["dx"]  # m
    '''
    The arrays rho and dx have unit information attached, which is carried
                                            over during any mathematical
                                            operation.
    Operations involving incompatible units are automatically detected and
                                            prohibited.
    '''
    mass = rho * (dx**3)  # kg
    return mass  # the exact units returned here do not matter, as the code
                                             uses dimensional analysis before
                                             returning the data to ensure SI
                                            units.
```

The method annotation takes a single argument with a list of the base fields (non-derived and known to RAMSES), such as gas density and cell width shown here, which are required to derive the desired field. Secondly, the function name must begin with 'amr', 'star' or 'dm', to denote whether this is a gas, star or dark matter derived quantity, followed by the field name (mass in this case). This function must always take the context and dset arguments, which correspond to the simulation snapshot and data dictionary which is guaranteed to always contain the required fields with unit information. Finally, the user simply implements the logic to derive the field and return the resulting array.

SEREN3 will ensure that the units are correct, and return the data to the user. Below, we show an example work flow which begins with loading a simple simulation object, and shows the steps to loading a series of derived and non-derived fields for an individual halo from a ROCKSTAR catalogue (Behroozi et al., 2013b). The data is automatically filtered to return particles and cells within the halo virial sphere only, from which we show how to calculate the halo baryon fraction below with ease. The arrays returned also support unit conversion, as shown.

```
import seren3

simulation = seren3.init("/path/to/simulation/directory")  # load a basic
                                        simulation object
iout_at_z = simulation.redshift(6)  # determine the simulation output
                                        nearest to z=6

snapshot = simulation[iout_at_z]  # load the snapshot using the simulations
                                         __getitem__ method
halos = snapshot.halos(finder="rockstar")  # load the rockstar halo
                                        catalogue. AHF is also supported,
                                        along with Merger Trees produced
                                        using ConsistentTrees
sorted_halos = halos.sort("Mvir")  # sort the halo catalogue based on
                                        virial mass

halo = sorted_halos[0]  # get the most massive halo

'''
Load some basic quantities for this halo. The computational domain is
                                        decomposed and data is filtered on
                                        the fly to return particles/cells
                                        within the halo virial radius only.
                                        OpenMP threading is enabled by
                                        default when loading data spanning
                                        multiple CPU domains.
'''
dm_dset = halo.d["mass"].flatten()  # load the dark matter halo mass.
gas_dset = halo.g[["rho", "mass", "T", "Gamma"]].flatten()  # load gas
                                        density, mass (derived), temperature
                                        (derived), and photoionization rate (
                                        derived)
star_dset = halo.s[["mass", "age"]].flatten()  # load star particle mass
                                        and age (derived from super-comoving
```

```
                                        coordinates to Gyr)

'''
Compute the baryon fraction for this halo. Unit tracking is built into the
                                        analysis framework, so returned
                                        arrays will always have
appropriate units attached. The resulting array also supports unit
                                        conversion as shown.
'''
dm_mass = dm_dset["mass"].in_units("Msol").sum()
star_mass = star_dset["mass"].in_units("Msol").sum()
gas_mass = gas_dset["mass"].in_units("Msol").sum()

total_mass = dm_mass + star_mass + gas_mass
baryon_fraction = (gas_mass + star_mass) / total_mass

'''
The Halo object can return basic information about each halo, such as ID,
                                        mass and virial radius, using the
                                        __getitem__ method.
'''
print "The baryon fraction of halo %s with mass %1.2e is: %f" % (halo["id"]
                                        , halo["Mvir"], baryon_fraction)
```

Chapter 3

Using Artificial Neural Networks to Constrain the Halo Baryon Fraction during Reionization

In this chapter, we investigate the self regulation of the halo baryon fraction as a result of the growing UV background during reionization. This is an important quantity in semi-analytical models of galaxy formation, with existing models relying on a simple fitting function based solely on mass and redshift. We go further, developing an artificial neural network which is trained using our self-consistent radiation-hydrodynamical simulations of galaxies during the EoR. The input parameters limited to purely physical quantities of haloes in our simulations, and we demonstrate that our network is independent of reionization history and redshift. Much of the content in this chapter can be found in Sullivan et al. (2017), which has been accepted for publication in MNRAS *as is*.

3.1 Introduction

During the first billion years after the Big Bang, the large-scale cosmic web of structures we see today began to form, followed by the first stars and galaxies, which brought an end to the Dark Ages (Rees, 1999). These first luminous sources are thought to be prime candidates to drive cosmic reionization, a major phase transition of the Universe, from a neutral intergalactic medium (IGM) following recombination, to the ionized state it still remains in today. This period is known as the Epoch of Reionization (EoR), and the wide range of physical processes which drive it encapsulate several areas of research, from cosmology and galaxy formation to radiative transfer (RT) and atomic physics. Even with

the wealth of present-day observational information at our disposal, these processes are not fully understood, and still less so in the high-redshift Universe.

Photoheating of the IGM by stars (and likely also by quasars and other radiation sources) raises the fully-ionized gas temperature to $T_{\rm IGM} \sim 10^4$ K, preventing the cooling and adiabatic collapse of gas into low-mass dark matter (DM) haloes. This quenching was originally discussed by Doroshkevich et al. (1967) and investigated by Couchman and Rees (1986) within the cold dark matter (CDM) framework, and may explain the discrepancy between the observed faint end galaxy luminosity function and the prediction from CDM, known as the missing satellites problem (Moore et al., 1999). Furthermore, the smallest dwarf galaxies can be photoevaporated entirely: three dimensional RT simulations that probe the propagation of ionization fronts (I-fronts) show that as supersonic R-type fonts encounter dense neutral gas in (mini) haloes, they quickly decelerate to D-type fronts, preceded by shock waves which blow the gas back into the IGM as an ionized supersonic wind (Shapiro et al., 2004; Iliev et al., 2005b). This occurs on time-scales of hundreds of millions of years, reducing the baryon fraction to a few percent (or less) of the total halo mass. As a consequence, the resulting increased photon consumption rate slows the propagation of global ionization fronts and extends reionization (Iliev et al., 2005a).

Recent studies have attempted to quantify the reduction of baryons in low-mass haloes using semi-analytical-models (SAMs) (e.g Babul and Rees, 1992; Efstathiou, 1992; Shapiro et al., 1994; Nagashima et al., 1999; Benson et al., 2002b,a; Somerville, 2002) and full 3D cosmological simulations (e.g Quinn et al., 1996; Weinberg et al., 1997; Gnedin, 2000a; Hoeft et al., 2006; Okamoto et al., 2008; Noh and McQuinn, 2014). In the absence of a photoionizing background, the baryon fraction $f_{\rm b} = M_{\rm b}/M_{\rm tot}$ of haloes scatters around the cosmic mean $\langle f_{\rm b} \rangle = \Omega_{\rm b}/\Omega_{\rm M}$. Supernovae (SNe) are not expected to significantly effect $f_{\rm b}$, as even the most efficient scenario (Sedov blast wave) can only accelerate gas to a few hundred kms^{-1}, while the escape velocity of a Milky Way sized halo is $v_{\rm esc} \sim 700$ km/s (and higher at high-redshift, where haloes are denser and more compact). As a result, SNe feedback is largely important *only* for dwarf galaxies (Dubois and Teyssier, 2008; Roškar et al., 2014), however, such events become increasingly rare towards lower masses. For a single stellar population, the available radiative energy of massive stars outweighs the cumulative feedback budget from galactic winds and SNe Type II explosions by a factor of ~ 100 (Roškar et al., 2014). While this depends on the interstellar medium (ISM) being able to efficiently absorb radiation, it suggests that radiative feedback could be a crucial ingredient in the regulation of star formation in galaxies.

Currently, SAMs rely on hydrodynamical simulations to constrain $f_{\rm b}$ in their galaxy formation models. Typically, one matches the halo mass with a baryon fraction half the cosmic mean, known as the *Characteristic mass*, $M_{\rm c}$. This scale sets the mid-point in the transition from baryon poor to baryon rich, thus can be used to generate simple fitting functions, which tightly constrain $f_{\rm b}$ as a function of halo mass. Prior to this work, $M_{\rm c}$ was thought to be set by the Jeans mass (or the *filtering scale*, which takes into account the halo mass assembly; Shapiro et al. 1994; Gnedin and Ostriker 1997; Gnedin 2000b), evaluated at the mean density of the Universe. The filtering scale was found to overshoot $M_{\rm c}$ during the EoR by Hoeft et al. (2006) and Okamoto et al. (2008), who instead argued that gas can accrete only if the haloes equilibrium temperature (where photoheating balances cooling), evaluated at over-densities in the range $60 \leq \Delta \leq 1000$, is greater than the virial temperature. However, even these studies were found to overshoot $M_{\rm c}$, as the true bottleneck for collapse occurs at densities that are an order of magnitude lower than the virial density and an order of magnitude above the cosmic mean density. At these densities, gas is not yet able to cool efficiently, while once it reaches the virial density, it will almost certainly continue to collapse as it can radiate away its energy efficiently (Noh and McQuinn, 2014). Clearly, there is a great deal of uncertainty surrounding the physical processes which set $M_{\rm c}$.

While computational algorithms have significantly improved since these studies, little advancement has been made on constraining the halo baryon fraction. Therefore, we are motived to revisit this in the context of fully self-consistent radiation hydrodynamical (RHD) simulations. We compare our results to the predictions of Okamoto et al. (2008) and Hoeft et al. (2006), developing and training an artificial neural network (ANN), which reproduces our simulations well. This model is independent of reionization history and can be incorporated into SAMs to self-consistently compute $f_{\rm b}$, better capturing the scatter.

This Chapter is arranged as follows. The details of our methods, star formation calibration, and SNe feedback scheme are outlined in Section 4.2. Our simulation set-up, along with key parameters and choice of halo finder is discussed in Section 3.3, where we also study radiative feedback on galaxies using a statistical analysis of haloes. In this section we also introduce our results for the halo baryon fraction for three different reionization histories. We describe our ANN architecture in Section 3.4, along with our choice of back-propagation algorithm and adaptive learning rate. We conclude this section by presenting our constraints on $M_{\rm c}(z)$, and use our ANN to predict this quantity based on our simulation results and a test model where we remove strong tidal forces. Finally, we discuss our

results in the context of the literature in Section 3.5 and present our final conclusions in Section 3.6.

3.2 Method

3.2.1 Radiation Hydrodynamics Code

To investigate the impact of feedback during the EoR in a cosmological context, we use the Eulerian adaptive mesh refinement (AMR) code RAMSES (Teyssier (2002); ver. 3). Basic grid elements are known as *octs*, which each belong to a particular level in the AMR hierarchy, ℓ, and are comprised of $2^{\rm ndim}$ cells, where ndim denotes the number of spatial dimensions (3 here). The code is based on the "Fully Threaded Tree" structure of Khokhlov (1998), where refinement is performed on a cell-by-cell basis following the user desired criteria. The hydrodynamics solver employs a second-order Godunov scheme, proven to capture shocks and accurately follow the thermal history of the fluid with hydrodynamical states reconstructed on cell interfaces using a MinMod method and advanced using the Harten-Lax-van-Leer contact wave Riemann solver (HLLC; Toro et al. 1994). The Poisson equation is solved using an adaptive particle-mesh method, with the collisionless N-body system described by the Vlasov-Poisson equations. Gas is allowed to cool through atomic excitation/de-excitation down to 10^4 K and further to ~ 100 K via metal fine-line transitions

We use the multi-group RT module RAMSES-RT, developed by Rosdahl et al. (2013), to compute the photoionization of three species: H I, He I, and He II. The code uses a moment-based scheme to solve the RT equation for three photon groups (one for each species) using a first-order Godunov method with the M1 closure for the Eddington tensor (Aubert and Teyssier, 2008, 2010). By treating the photons as a fluid in this way, the code makes natural use of the existing solvers implemented in RAMSES to fully couple the radiation with the gas via photoionization and photoheating and a set of non-equilibrium chemistry equations for the three species above. This scheme has the advantage over ray-tracing methods that the RT step is independent of the number of sources, ideal for cosmological simulations of galaxy formation. The production rate of ionizing photons varies with time for a given IMF, with the majority of photons released within 5 Myr from the birth of a star particle. We adopt the *on-the-spot* approximation, where UV photons emitted from recombinations are assumed ionize an atom in the same grid cell, hence case-B recombination rates are used when computing the gas cooling rate (Hui and

Gnedin, 1997).

3.3 Cosmological Simulations

We assume a geometrically flat ΛCDM Universe consistent with a WMAP-7 cosmology (Komatsu et al., 2011) with cosmological parameters: $\Omega_{\rm M} = 0.272$, $\Omega_\Lambda = \Lambda_0/(3H_0^2) = 0.728$, $\Omega_{\rm b} = 0.045$, $h \equiv H_0/(100 \text{ km s}^{-1} \text{ Mpc}^{-1}) = 0.702$, and $\sigma_8 = 0.82$, where the symbols take their usual meaning. Initial conditions are generated using the MUSIC code (Hahn and Abel, 2013) at redshift $z = 100$ in a comoving volume of ~ 5.7 Mpc on-a-side. Variations in these cosmological parameters are not expected to largely influence our results.

The coarse grid level is set to $\ell_{\rm min} = 8$ in the AMR hierarchy for all simulations, resulting in $(2^{\ell_{\rm min}})^3 = (256)^3$ DM particles and coarse grid cells, such that our DM particle mass $M_{\rm DM} \sim 3.4 \times 10^5$ M$_\odot$. We set the maximum level of refinement $\ell_{\rm max} = 18$, such that our effective spatial resolution is $\Delta x_{\rm min} = 21.72$ pc comoving and our stellar particle mass is $m_{*,\rm min} \sim 2 \times 10^3 M_\odot$. We adopt uni-grid initial conditions for the gas and dark matter (i.e.., no zoom regions).

We calibrate our simulations and adopt star formation and SNe feedback parameters based on the methodology outlined in sections 2.7.1 and 2.7.2, with parameters:

$$\begin{aligned} n_* &\geq 8 \text{ cm}^{-3} \\ T_0 &= 540 \text{ K}, \\ g_* &= 2, \end{aligned} \tag{3.1}$$

where g_* is the polytropic index. This ensures the Jeans length of star-forming gas is always resolved by at least six cells in our simulations.

Dark matter (sub)-haloes are catalogued using the ROCKSTAR phase-space halo finder (Behroozi et al., 2013b), while merger trees are constructed using the CONSISTENT-TREES algorithm (Behroozi et al., 2013c). Particles are first divided into 3D Friends-of-Friends (FOF) groups, which are analysed in 6D phase-space to give robust, grid- and shape-independent haloes. Haloes that are not gravitationally bound are removed and properties are computed using the virial overdensity definition of Bryan and Norman (1998).

Photoionization of both hydrogen and helium is followed via the coupling of UV photons with the gas, assuming a Bruzual and Charlot (2003) stellar population for an instantaneous Chabrier IMF (Chabrier, 2003). We perform a suite of RHD simulations

Name	Radiative Transfer	$f^*_{\rm esc}$
RT1.5	Yes	1.5
RT2 (Fiducial)	Yes	2.0
RT5	Yes	5.0
HD	No	N/A

Table 3.1: List of simulations. We denote all RT simulations as RTX, where X refers to the choice of stellar escape fraction, $f^*_{\rm esc}$. For reference, we include a model with RT disabled (i.e.., SNe feedback only), which we denote HD. We choose RT2 as our fiducial model, as its reionization history falls between our other two models (see figure 3.2).

with stellar particle escape fractions, $f^*_{\rm esc} = 1.5,\ 2$, and 5, which is multiplied by the total number of photons integrated from our spectral energy distribution (SED) when emitting from star particles as listed in Table 3.1. As the star formation model is sub-grid, this factor accounts for additional uncertainties such as the increases efficiency of Population III stars, therefore could in principle exceed unity. Due to computational constraints, we adopt a reduced speed of light, $f_{\rm c} = 1/100$, which relaxes the Courant-condition for the radiation fluid and, thus, reduces the number of radiation sub-cycles necessary to reach thermal convergence (Gnedin and Abel, 2001). We choose simulation RT2 as our fiducial model, as its reionization history falls between our RT1.5 and RT5 models.

Several authors have investigated the evolution and geometry of the EoR on large-scales, which accurately capture the statistics of this process (e.g Ciardi et al., 2003a; Iliev et al., 2006b; Mellema et al., 2006b; Zahn et al., 2007; McQuinn et al., 2007; Iliev et al., 2007a; Mesinger and Furlanetto, 2007; Geil and Wyithe, 2008; Choudhury et al., 2009; Dixon et al., 2016), and the general consensus is that reionization progresses inside-out, with high-density regions being ionized earlier, on average, than the voids. In such a scenario, the complex interplay between the growth of H II regions and the formation of structure in the Universe is crucial; while I-fronts from the first sources can quickly escape into voids, the exponential growth of structure during this extended process ensures that there are always new stars forming in the high-density peaks, which overwhelm reionization. The mass-weighted neutral fraction therefore remains steadily lower than the volume-weighted quantity throughout. To properly follow the inside-out progression of reionization, large-volumes are required which fully capture the statistics of structure formation (Mellema et al., 2006a; Iliev et al., 2006b, 2014).

Early small-volume simulations predict an outside-in progression (e.g Gnedin 2000a), with I-fronts quickly escaping into voids and slowly eating their way through filaments. Due to the limited volume, these models do not form as many new sources capable of overwhelming reionization; therefore, the opposite effect is found, whereby the volume-weighted neutral fraction lags behind the mass-weighted value. In figure 3.2, we show the evolution of both the volume-weighted and mass-weighted (solid and dashed, respectively) neutral fraction (top panel), photoionization rate (Γ; middle panel), and the integrated electron-scattering optical depth (τ; bottom panel) for each of our RT models, along with relevant constraints on each quantity.

For all models the mass-weighted neutral fraction is consistently below the volume-weighted quantity during the EoR, which is characteristic of an inside-out progression. This is depicted in figure 3.1, where we show mass-weighted projections of the fraction of ionized hydrogen, $x_{\rm HII}$, of the full volume at redshifts $z = 12$, 9, 8.5 and 7.5. The initial growth of H II regions stems from the high density peaks (top left panel), and gradually grow before percolating and the completion of hydrogen reionization. This occurs shortly after $z = 7.5$, where in the bottom right panel we can still see a few pockets of neutral gas in underdense regions. Ionization fronts sweep through these regions last, leading to a steep drop in the volume-weighted neutral fraction at the end of the EoR as they make up a large fraction of the volume (in contrast to the mass-weighted value, as the solid lines converge/overtake the dashed lines).

Our fiducial model shows good agreement with constraints on Γ and τ, however is in slight tension with neutral fraction constraints from Lyman-α emitters. While this tension is not entirely relieved by our RT1.5 model, it shows much better agreement for both the neutral fraction and PlanckTT+lowP+lensing+BAO 2016 constraints on τ (Planck Collaboration et al., 2016a). As expected, our early reionization model, RT5, is in tension with all constraints with reionization ending at $z_{\rm reion} \sim 8.5$, however we note that it is consistent with the PlanckTT+lowP+lensing+BAO 2015 constraints on the optical depth within 1σ ($\tau = 0.066 \pm 0.013$, Planck Collaboration et al. 2016b). We include this model to test the sensitivity of our results on the desired reionization history.

The difference between the volume- and mass-weighted neutral fraction becomes smaller as we increase the emissivity of stellar particles, as the increased production rate of photons boosts the ability of new sources to dominate the EoR. In our RT1.5 and RT2 models, the mass-weighted fraction levels out above zero, while the volume-weighted fractions of all simulations converge on zero, as the combination of an increased recombination

rate and decreased stellar luminosity (lower emissivity) allows high-density regions to recombine on shorter time-scales.

The mass-weighted photoionization rates are significantly above the volume-weighted averages and constraints from Calverley et al. (2011) and Wyithe and Bolton (2011), as ionization fronts emanate from high-density peaks and sweep through underdense regions much later. Due to the limited volume/statistics of our box, both the volume and mass-weighted photoionization rates show significant fluctuations at high-redshift. Towards the end of the EoR, the volume-weighted quantity smooths out as voids become ionized. The mass-weighted continues to show some fluctuation, especially in our high emissivity case (RT5), due to the balancing of photoionization and recombinations in high-density regions.

3.3.1 Suppression Mechanisms

The significance of self-feedback by radiative heating/ionization is heavily debated in the literature, from its magnitude to whether it is a factor at all (see, for e.g Shapiro et al. 1994; Gnedin 2000b; Shapiro et al. 2004; Iliev et al. 2005b; Hoeft et al. 2006; Wyithe and Loeb 2006; Okamoto et al. 2008; Kimm and Cen 2014; Ocvirk et al. 2016). Simulations of isolated disk galaxies show that radiative feedback acts primarily to prevent the adiabatic collapse of gas to higher (star-forming) densities, rather than push the gas past some temperature threshold deemed important for star formation (Rosdahl et al., 2015). We repeat this investigation in a cosmological context, using our HD simulation to understand the impact of including radiative feedback in our RT models, focusing on the temperature and density distribution of gas in DM haloes.

The evolution of gas in the temperature-density plane is shown in figure 3.3 for our HD (left) and fiducial (right) models, where the vertical/horizontal dashed lines show the density and temperature criteria for star formation respectively. In the absence of radiative heating, the IGM remains cold with most of the gas occupying the adiabat $T/\mu \propto \rho^{\gamma-1}$, where γ is the ratio of specific heats. Adiabatic collapse is therefore uninhibited; while in our fiducial model, photoheating makes the gas nearly isothermal at $T/\mu \sim 10^4$ K. The two visible branches in the right panel at densities $n_{\rm H} \gtrsim 10^3 {\rm m}^{-3}$ result from atomic cooling peaks of hydrogen (lower) and helium (upper), which is highly efficient at these densities. This process allows the gas to collapse to the metal line cooling regime, reaching $T/\mu \sim 10^3$ K/μ at densities $n_{\rm H} \gtrsim 10^4 {\rm m}^{-3}$. Following the completion of the EoR (overlap stage), regions which are ionized later than average, predominately diffuse gas in the circumgalactic medium (CGM) or voids, remain hotter on average due to their longer

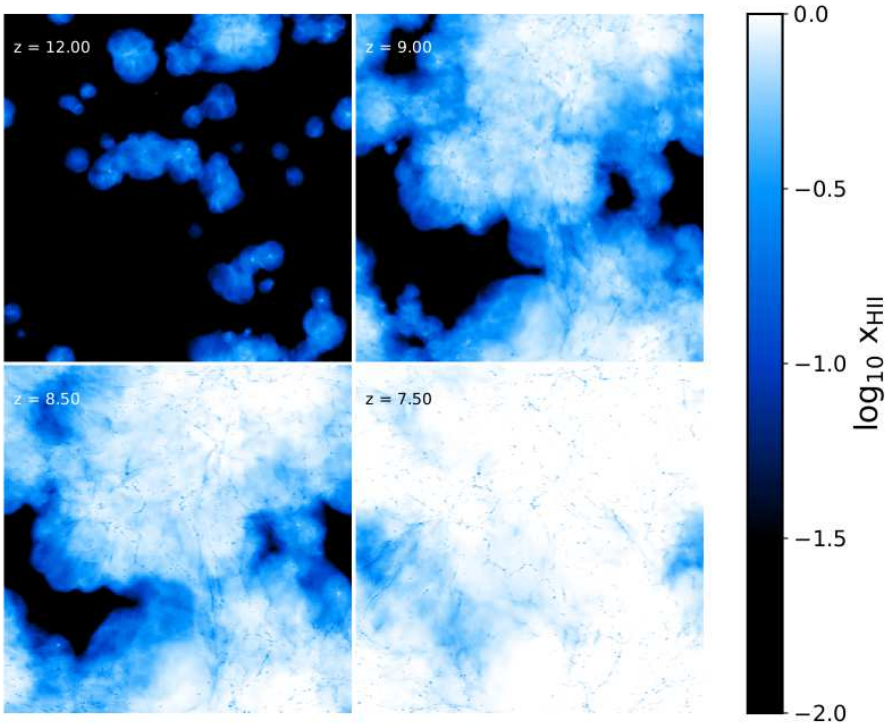

Figure 3.1: Mass-weighted projections of the hydrogen ionization fraction, H II, at redshifts 12, 9, 8.5, and 7.5 in our fiducial RT2 model. Each panel shows the full projection of the $(5.7 \text{ Mpc})^3$ volume, projected along the z-axis. At $z = 12$, we see the formation of H II regions around the first stars in the universe, which grow and eventually overlap completing the EoR. Our model exhibits an inside-out progression, consistent with the literature and discussed in the text, whereby voids are the last regions to undergo reionization (as seen in the final panel; $z = 7.5$).

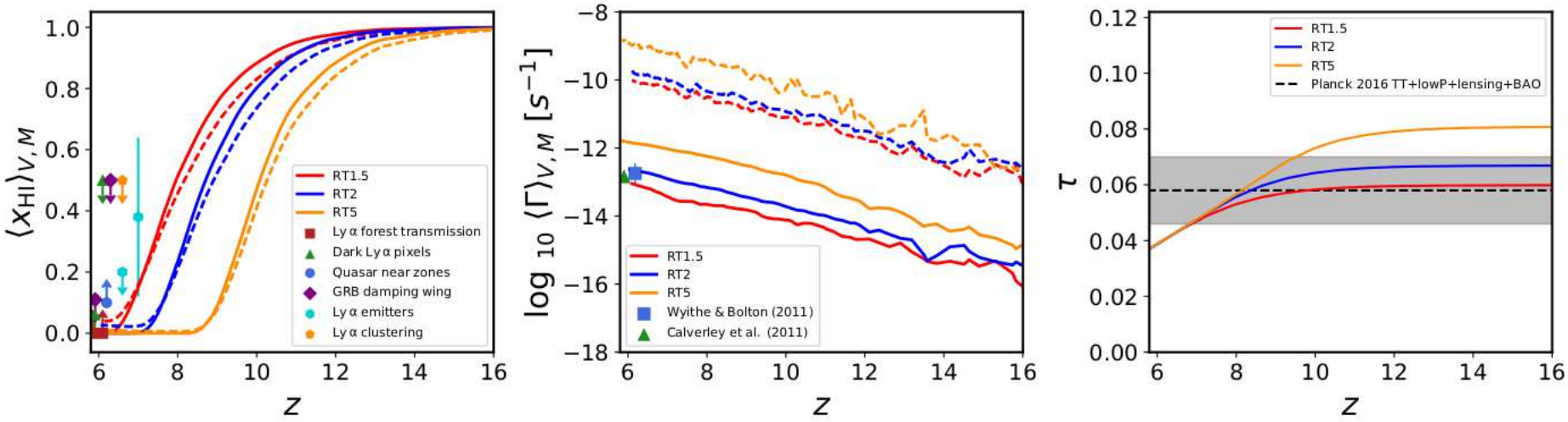

Figure 3.2: *Top panel:* the volume/mass-weighted (solid/dahsed respectively) mean neutral fraction of hydrogen for all RTX simulations listed in Table 3.1, compared to observational constraints from Ly α forest transmission (squares, red; Fan et al., 2006), dark Ly α forest pixels (triangles, green; McGreer et al., 2011, 2015), quasar near zones (circles, blue; Schroeder et al., 2013), GRB damping wind absorption (diamonds, voilet; McQuinn et al., 2008; Chornock et al., 2013), and decline in Ly α emitters (hexagons, cyan; Ota et al., 2008; Ouchi et al., 2010) following from Robertson et al. (2015). *Middle panel:* The mean volume/mass-weighted photoionization rate, compared to the volume-weighted observations of Calverley et al. (2011) and Wyithe and Bolton (2011) as the (green) triangle and (blue) square, respectively. *Bottom panel:* The integrated electron-scattering optical depth compared to the PlanckTT+lowP+lensing+BAO 2016 results (thin black horizontal line) and the 1σ error interval (shaded region; Planck Collaboration et al., 2016a).

cooling time ($t_{\rm cool} \sim 1/n_{\rm H}$) and additional local heating due to shocks from SNe and structure formation.

In both cases, feedback from SNe heats the gas to temperatures 2×10^4 K $< T/\mu <$ 10^7 K, with noticeably less SNe heating when RT is included. As we will show below, this lack of SNe shock heating is due to the additional suppression of star formation by radiative feedback. As such events are already rare, any additional modulation of the star formation rate will have a significant impact on the SNe feedback energy budget. To determine whether radiation primarily acts to heat the gas above the star formation criteria or prevent collapse in the first place, we compute the cumulative distribution function (CDF) of densities and temperatures (left/right respectively) in haloes, binned by mass, as shown in figure 3.4.

As gas infalls from the IGM onto haloes, it undergoes a strong virial shock, which heats it to of order the virial temperature. This heating leads to a hot diffuse atmosphere, which occupies the CGM. Irradiation can keep the gas in this diffuse state for prolonged periods of time, by extending the cooling time and thus preventing collapse. This behaviour is clearly demonstrated for our fiducial RT2 simulation in the bottom left panel of figure 3.4 (solid lines), when compared to our SNe only case (dashed lines), where we show the difference between both cases in the top panel. Radiative feedback significantly alters the statistical distribution of gas densities in haloes of all masses, with low-mass haloes most strongly affected (due to their shallow potential wells). This feedback leads to a reduction in gas density several orders of magnitude below n_* (vertical dot-dashed line).

Interestingly, we find haloes in the highest mass bin are, on average, more suppressed in the high density regions than those in the $10^9\,M_\odot h^{-1} \leq {\rm M_{vir}} < 10^{10}\ M_\odot h^{-1}$ bin when RT is included. This additional suppression is due to the amplification of SNe by radiative feedback and vice versa, as first discussed in Pawlik and Schaye (2009). As the gas is kept in a diffuse state, a Sedov blast wave can much more efficiently blow gas from the centre of the halo to the outer regions, where in turn it is more susceptible to photevaporation (Pawlik and Schaye, 2009). While this phenomenon is present in all haloes, it is most effective in high-mass haloes, where the frequency of SNe is higher.

In the right-hand panel, we compute the same quantity but for the gas temperature, where we have subtracted the Jeans pressure floor introduced in Section 2.7.2. The distribution of temperatures for all halo mass bins are roughly equal in our RT case, as photoheating raises the gas temperature to ~ 8000 K. In the absence of radiative heating, low-mass haloes contain significantly more cold gas due to their weak virial shocks and rare

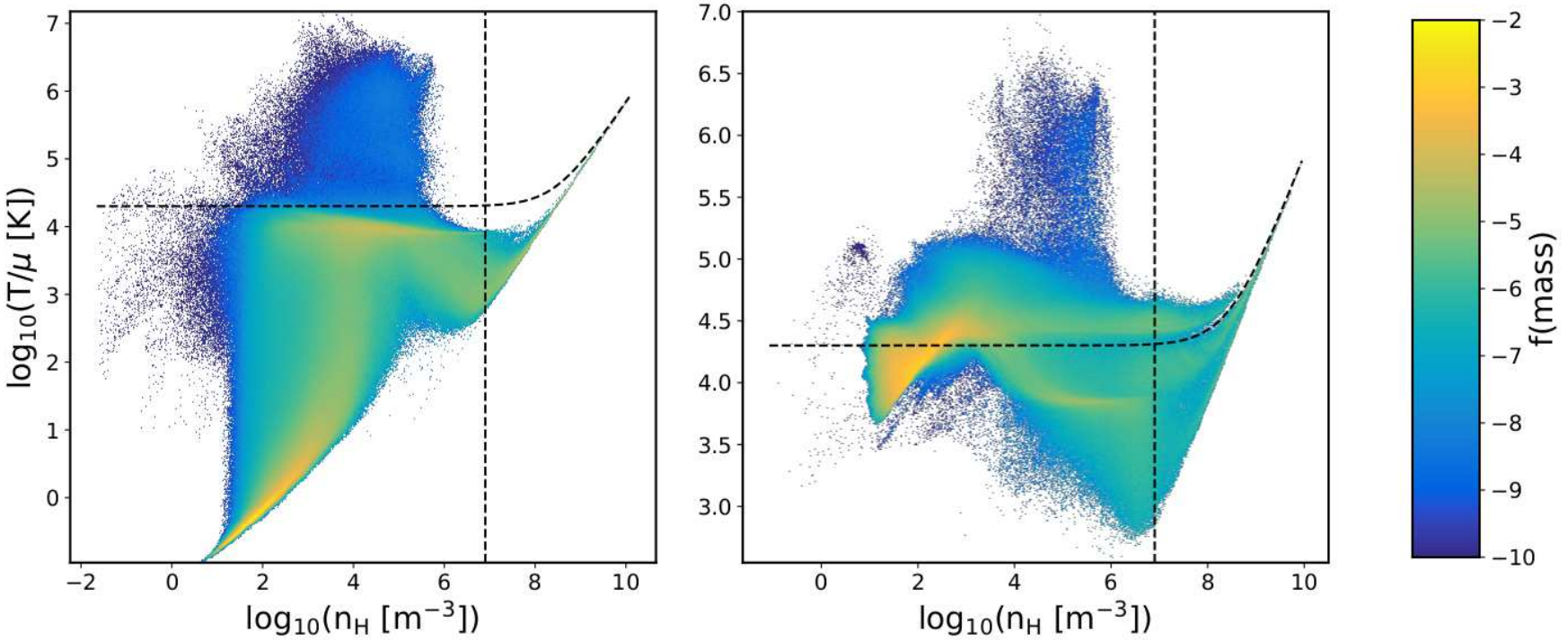

Figure 3.3: Mass-weighted temperature-density phase diagrams for our HD and RT2 simulations (left/right respectively), for the full box at $z = 6.15$. The dashed lines represent the density and temperature criterion for star formation (vertical/horizontal), as discussed in Section 2.7.2, which bound the star-forming region of the plot (above/below the density/temperature criterion, respectively). The change in shape of the temperature criterion at $n_{\rm H} = n_*$ corresponds to the point where our polytropic temperature floor begins to dominate. In the SNe only case (HD), we see that most of the gas by fraction of mass remains on the adiabat $T/\mu \propto \rho^{\gamma-1}$, where γ is the ratio of specific heats. In the absence of a UV background, the gas is allowed to continue with adiabatic collapse into haloes, leading to run-away star and galaxy formation. In the RT2 case, the EoR raises the gas pressure of the IGM, preventing adiabatic collapse and raising the Jeans mass of galaxies. To collapse to higher densities, gas must cool along the visible atomic cooling branches at $T/\mu \sim 10^{4.6}$ K and $T/\mu \sim 10^{3.8}$ K, which corresponds to the helium and hydrogen cooling peak cut-offs respectively. Once the gas reaches high densities, metal fine-line cooling becomes dominant and allows the gas to cool to $T/\mu \sim 10^3$ K in both cases. At these high densities, gas is capable of self-shielding from ionizing UV, allowing star formation (see also Appendix 3.A).

SNe events, bringing the mean gas temperature several orders of magnitude below the star forming criterion (vertical dot-dashed line). However, in the context of whether a cell can become star forming or not, we are only concerned with the properties of the ISM, where the typical gas density is $n_{\rm H} \sim 1$ cm^{-3}. At these densities, cooling is extremely efficient as shown in figure 3.3 and therefore not highly sensitive to the temperate criterion for star formation. Therefore, the primary mechanism by which radiation inhibits star formation is to prevent the otherwise efficient collapse from the IGM into galaxies.

Figure 3.5 shows the mean star formation rate (SFR) in logarithmic bins of halo mass (taken at $z = 6$) for our fiducial and HD models (left/right respectively), from high-mass (red) to low-mass (grey), as a function of redshift. High-mass haloes show very little suppression due to radiative feedback, except at early times when their progenitors are affected. As a result, large haloes typically undergo internal (inside-out) reionization, while low-mass haloes on average are reionized by distant sources (outside-in). This pattern is supported by simulations of the Local Group of galaxies, which find that low-mass satellites preferentially undergo external reionization due to the inside-out reionization imprinted by massive haloes (Weinmann et al., 2007; Ocvirk et al., 2014). While we do find that the SFR is increasingly suppressed towards lower halo masses, low-mass dwarfs are able to continue forming stars throughout the EoR, in contrast with recent works with lower numerical resolution (e.g. Ocvirk et al., 2016).

Similarly, in figure 3.6, we show the (log) stellar abundance as a percentage of the total halo mass, in units of the cosmic mean baryon fraction, at $z = 6.15$ for all models. We compare our results with the abundance matching models of Moster et al. (2013) (grey) and Behroozi et al. (2013a) (blue), where the sold line/shaded region show the mean and 1σ standard deviation respectively, and the dashed lines show the extrapolation below their minimum halo mass, $M_{\rm vir} \leq 10^{10}\ M_{\odot}$. We note that the Behroozi et al. (2013a) constraints are consistent with the observed galaxy stellar mass functions specific star formation rates and cosmic star formation rates from $z = 0$ to 8; however the Moster et al. (2013) model only considers haloes from $z \sim 4$ to present, hence we extrapolate to our redshift range.

The reduction in stellar mass, as measured against our HD model (black), is proportional to the assumed sub-grid stellar emissivity, and thus the reionization history (red to orange, increasing). While we see a large reduction in the mean, only our RT5 model shows tension with the SNe-only case within the 1σ error bars at low-masses. All of our models are well within the extrapolated (dashed lines) constraints of Moster et al.

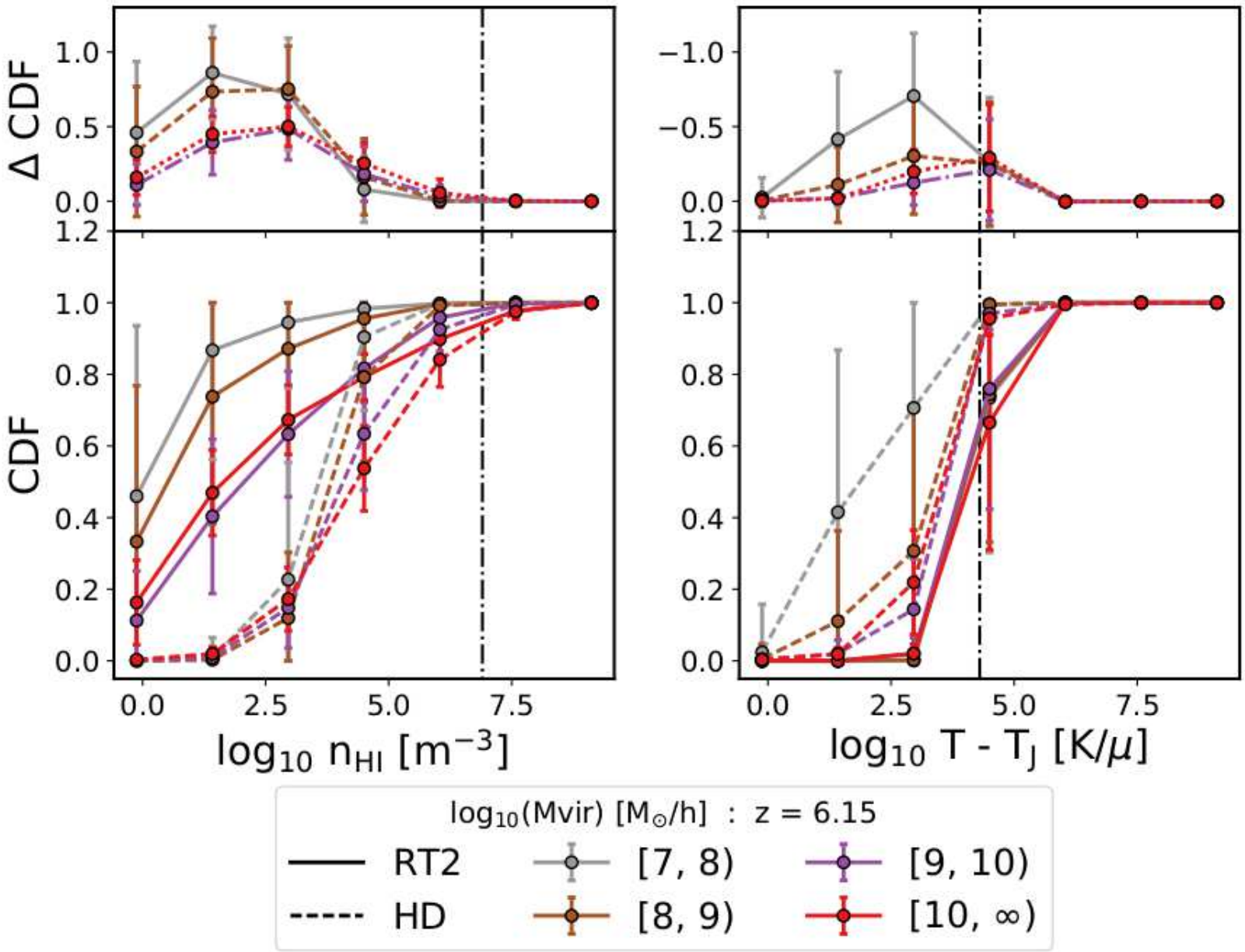

Figure 3.4: *Bottom Left panel:* Mean cumulative distribution functions (CDF) of neutral hydrogen number density in logarithmic bins of halo mass for our fiducial RT2 model (solid) and SNe only (dashed) HD model at $z = 6.15$, where the error bars show the 1σ standard deviation. The vertical dashed black (thin) line shows the density criterion for star formation, n_*. *Bottom right panel:* Mean CDF of gas temperature minus the

(2013), however radiative feedback is needed to bring our models into agreement with the extrapolated Behroozi et al. (2013a) constraints for intermediate halo masses.

3.3.2 The Halo Baryon Fraction

We compute $f_{\rm b}$ for each halo by summing the total mass of gas, stars and DM enclosed within the virial radius, $R_{\rm vir}$, as shown in figure 3.7. From left to right, we show the evolution over the redshift range $9.5 \lesssim z \lesssim 6$ for our fiducial model. The rows highlight the impact of four different quantities: the tidal force averaged over a dynamical time, $\langle F_{\rm tidal} \rangle_{t_{\rm dyn}}$, mass-weighted hydrogen ionization fraction, $\langle x_{\rm HII} \rangle_{\rm M}$, mass-weighted gas temperature in units of the virial temperature, $\langle T \rangle_{\rm M} / T_{\rm vir}$, and finally the (virial) ratio of kinetic to potential energies, $T/|U|$. We restrict our study to *distinct* haloes only, as including sub-haloes may artificially inflate $f_{\rm b}$.

Prior to the EoR (left-hand column), $f_{\rm b}$ scatters about the cosmic mean for high-mass haloes, $\rm M_{vir} > 10^8\ M_{\odot} h^{-1}$, and falls to roughly half this value at lower masses (black dashed line with error bars). In contrast, following the completion of reionization (centre and right-hand columns), low-mass haloes that are below the atomic cooling limit are quenched of their baryons, resulting in $f_{\rm b}$ falling to a few percent (or lower) of the total halo mass. In the high-mass regime, potential wells are deep enough to retain gas in the presence of an ionizing UV background, hence are not strongly effected. At all times, we find that a small percentage of the total halo population exhibit super-cosmic mean values of $f_{\rm b}$. Okamoto et al. (2008) found that these were subhaloes at earlier times that have been tidally stripped from their hosts, leaving a dense baryon core. Such phenomena are expected in a hierarchical ΛCDM model of galaxy formation, and in extreme cases the entire galaxy can be disrupted leaving only a stellar stream (as is thought for the Sagittarius dwarf galaxy; Belokurov et al., 2006). To quantify this effect, we consider the strongest tidal force experienced from all neighbours (figure 3.7, top panel). The tidal force is defined as the ratio of the virial radius and the radius of the Hill sphere:

$$R_{\rm Hill} \simeq d \left(\frac{m}{3M} \right)^{1/3}, \tag{3.2}$$

where d is the distance between the haloes, and m and M are the masses of the smallest and largest of the two haloes under consideration, respectively. To determine whether the halo has had time to adjust, we average the tidal force over a single dynamical time $t_{\rm dyn} = (4\pi G \rho)^{-1/2}$.

To better understand how the tidal force is distributed in the halo population, we show the probability (and cumulative) distribution functions (PDF and CDF respectively) for

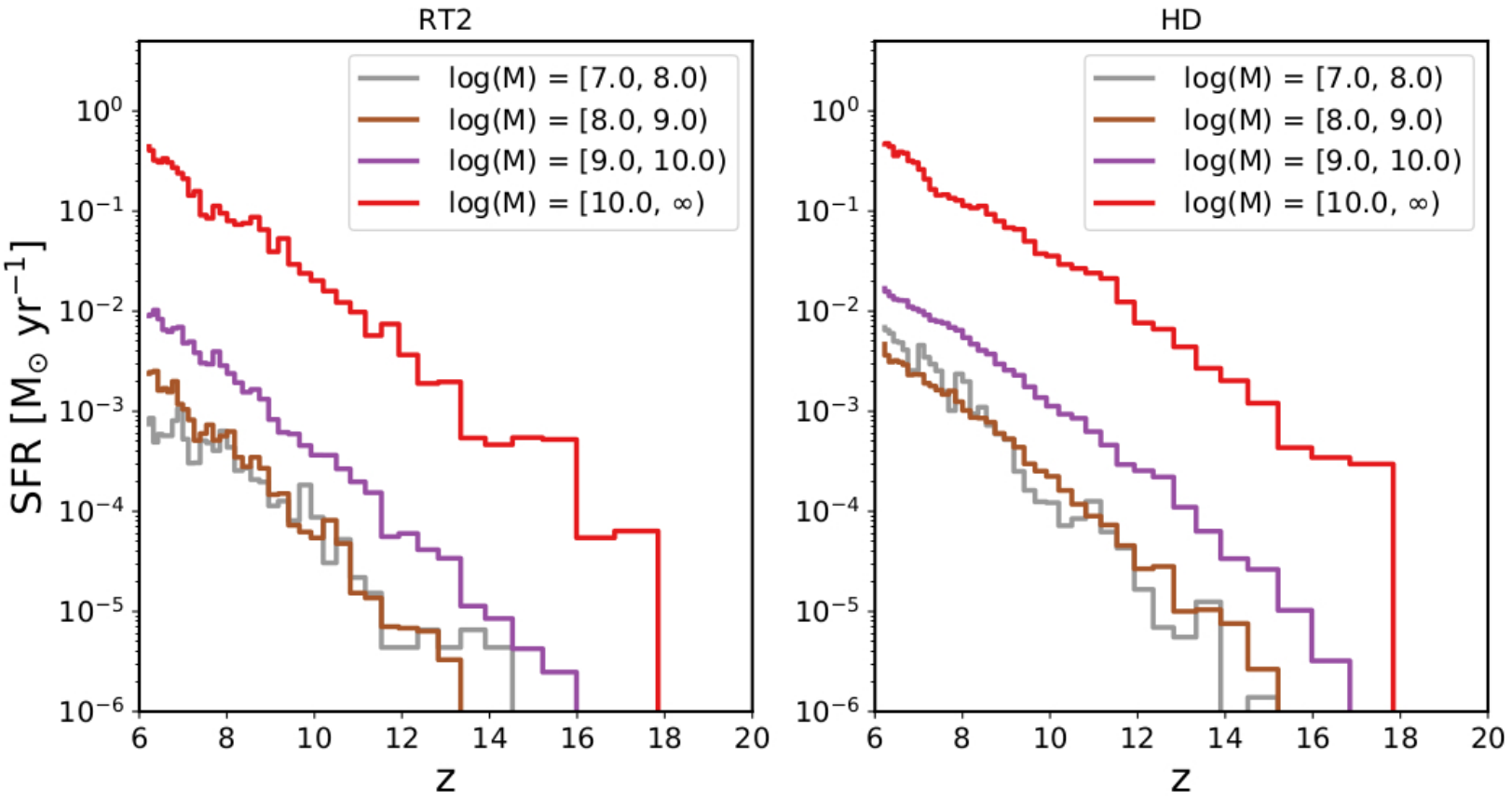

Figure 3.5: Star formation rates as a function of redshift in logarithmic bins of halo mass, in units of $M_{\odot}h^{-1}$, for our fiducial RT2 model (left) and HD (right). We compute the SFR by binning stellar particles in all haloes from $z = 6$ back to the formation of the first star. Radiative feedback increasingly suppresses star formation towards low halo masses (by comparing left to right). Massive haloes are not strongly effected except at early times, when their progenitors are suppressed. By $z = 6$, the SFR in the lowest mass halo bin has fallen by an order of magnitude with respect to the SNe only case, though does not cease entirely.

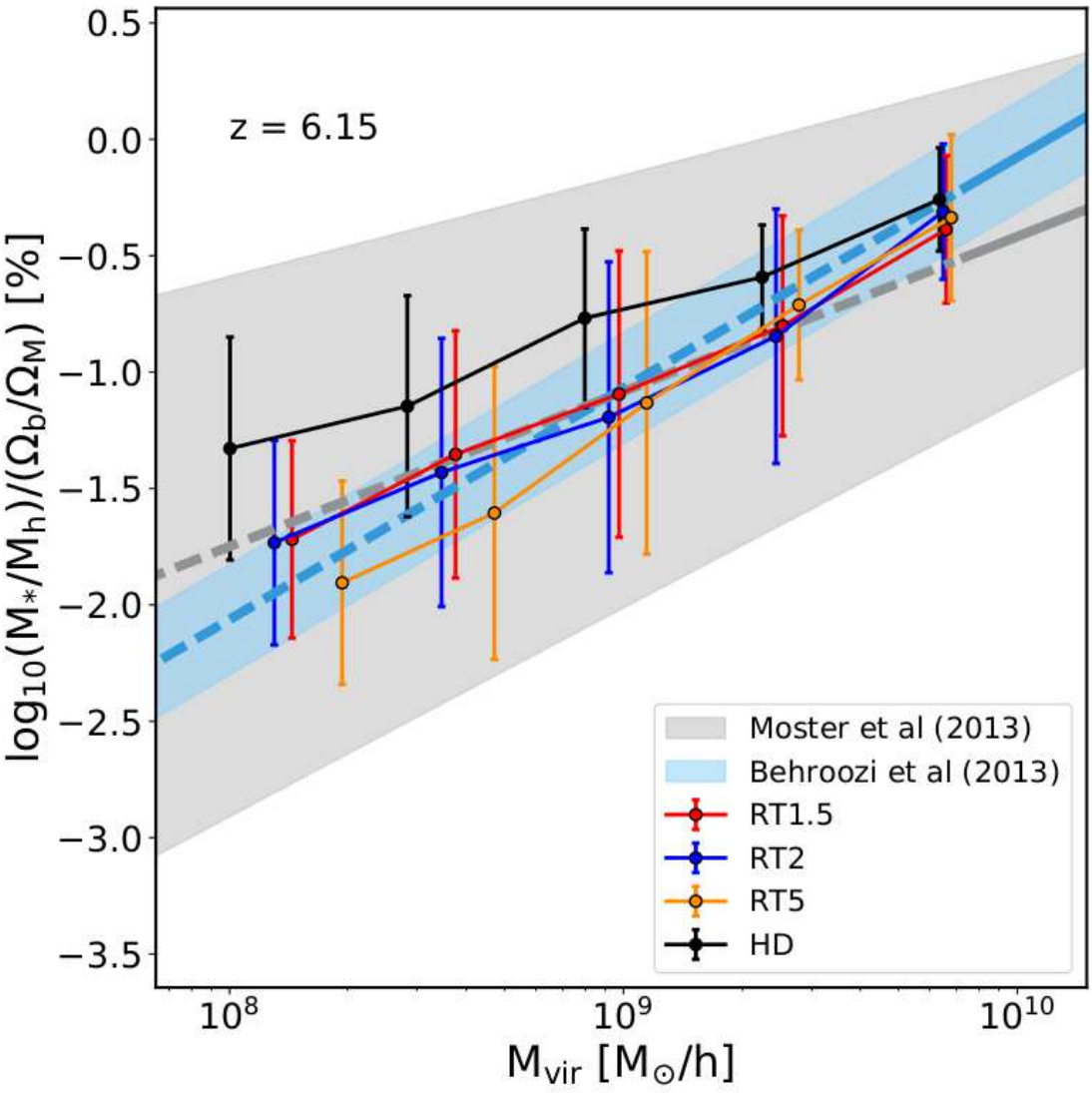

Figure 3.6: Stellar fraction of haloes (in units of the cosmic mean baryon fraction) for all simulations listed in Table 3.1 (points), compared with the models of Moster et al. (2013) and Behroozi et al. (2013a). Solid lines show the mean fit to the scatter, the error bars show the 1σ deviation, and the dashed lines show the extrapolated values for both models. While haloes with masses $\mathrm{M_{vir}} > 10^8\ M_\odot h^{-1}$ (i.e. $T_{\rm vir} \sim 10^4$ K) should not be strongly effected by UV heating, the legacy effect on their accumulated stellar mass, as their progenitors are suppressed, can be seen here when contrasted with the HD (SNe-only) case. Radiative feedback brings all RTX models into good agreement with the abundance matching model of Behroozi et al. (2013a) by $\mathrm{M_{vir}} \sim 8 \times 10^8\ M_\odot h^{-1}$, while all models fall well within the constraints of Moster et al. (2013).

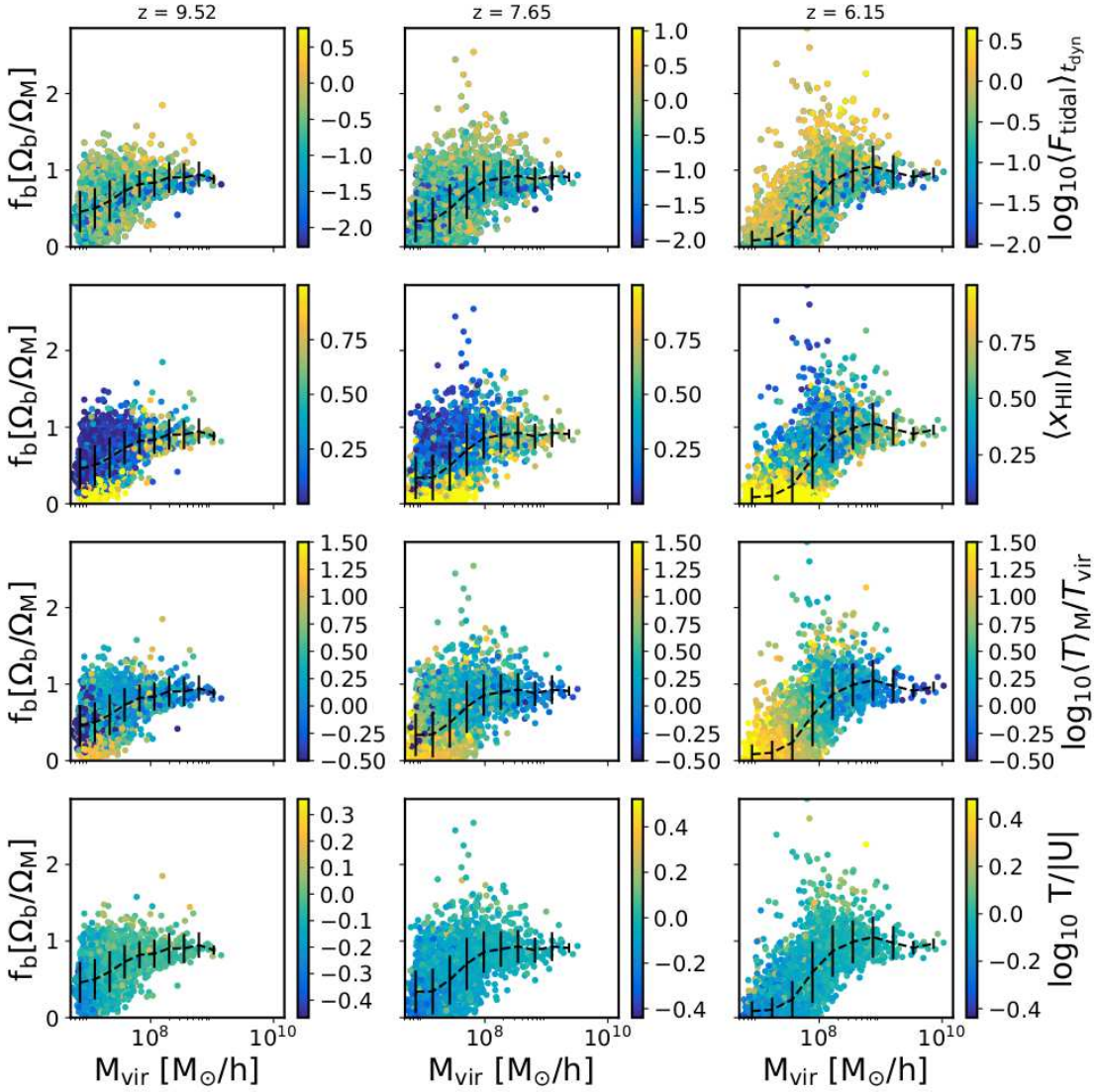

Figure 3.7: Halo baryon fraction, in units of the cosmic mean, at redshifts 9.52, 7.65, and 6.15 for our RT2 simulation (columns; left to right). The colour bars show (top to bottom) the strongest tidal force exerted over the last dynamical time, mass-weighted hydrogen ionization fraction, mass-weighted gas temperature (in units of the virial temperature) and the virial ratio (ratio of kinetic to potential energies). The dashed lines show the mean baryon fraction in bins of logarithmic halo mass, where the error bars correspond to the standard deviation.

the tidal force in figure 3.8 between redshifts $9.5 \geq z \geq 6$ for our RT2 fiducial model. The PDF exhibits a strong bimodal distribution at all times, where the small peak characterises isolated haloes (near or in voids/underdense regions) that have not undergone any tidal interactions. Over time, the amplitude of the smaller peak slowly decreases as isolated galaxies eventually interact and (possibly) merge with more massive hosts. Approximately half of all haloes undergo strong tidal disruption, by which we mean the tidal force exceeds the mean of the larger peak, as shown by the CDF.

By comparing with the top row of figure 3.7, we find that strong tidal disruption (yellow points) leads to a deviation above/below the mean in all mass bins (black dashed line). This deviation is characteristic of a hierarchical model of galaxy formation, as haloes with large baryon fractions have stripped material from their neighbours, which are left baryon poor. In contrast, haloes that have not undergone significant disruption in their last dynamical time (blue points) fall on or close to the mean. The fraction of haloes in isolation that undergo very weak or no tidal forces decreases from ~ 0.1 at $z = 9.52$ to < 0.05 at $z = 6.15$, as demonstrated by the drop in the CDF where it intersects the lower peak. This leads to a reduction in the density of points which lie along the mean $f_{\rm b}$ in each mass bin, and therefore an increased scatter as shown by the error bars. High tidal forces typically lead to above average baryon fractions over time (yellow points; left to right), suggesting that tidal stripping may play a crucial role during the EoR; enabling haloes that might otherwise be devoid of gas to continue accretion. If this gas can cool/self-shield and condense towards the centre of the halo, this would enable the prolonged growth of the central galaxy, and continued star formation.

In the second row of figure 3.7, we show the impact of photoionization (indicated by the ionized fraction, second row of plots) on $f_{\rm b}$. As the EoR progresses (left to right), the fraction of dwarf haloes that are completely ionized increases, bringing the mean $f_{\rm b}$ down to a few percent of the total halo mass. In the centre column, reionization has only just finished (see figure 3.2), thus there has not been sufficient time for the photoevaporation of all haloes below the atomic cooling limit ($\rm M_{vir} \sim 10^8\ M_{\odot} h^{-1}$).

We find that immediately following the EoR, there are still dwarf haloes that contain cold gas (third row; dark blue points). If this gas is sufficiently dense, it may self-shield from the ionizing UV background, extending the time required to photoevaporate the halo fully. Between $z = 7.65$ and $z = 6.15$ (centre and right column), approximately 230 Myr passes, in which time the gas temperature rises and the photoevaporation of dwarfs takes place, bringing the mean $f_{\rm b}$ down from roughly half the cosmic mean to a few percent of

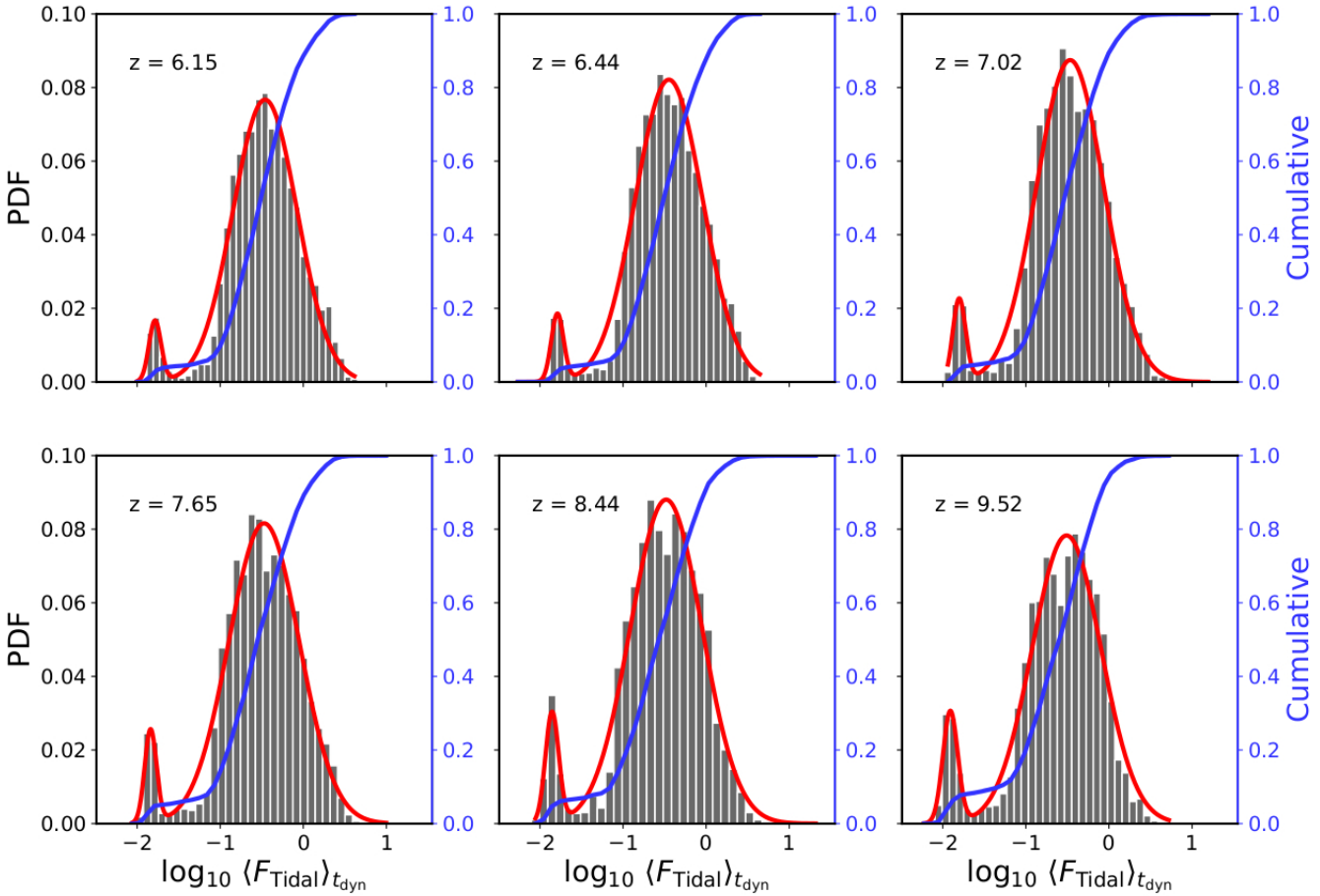

Figure 3.8: Probability distribution functions (PDFs) and Cumulative distribution functions (CDFs) for the tidal force averaged over a dynamical time, $\langle F_{\rm tidal} \rangle_{t_{\rm dyn}}$, from redshifts $6 \lesssim z \lesssim 9.5$. We fit a bimodal Gaussian distribution to each PDF (red solid lines). At all redshifts, approximately half of all haloes experience tidal forces less than or equal to the position of the larger peak, while the amplitude of the smaller peak slowly decreases over time as isolated galaxies eventually interact and (possibly) merge with more massive hosts.

the total halo mass. At intermediate masses, between the gas poor and gas rich extremes, haloes can cool via atomic de-excitation and collisional processes, allowing the gas to recombine.

Above these masses, $f_{\rm b}$ approaches the cosmic mean, despite large fractions of the gas being ionized (second row; yellow points at masses $M_{\rm vir} \gtrsim 10^8\ M_{\odot}h^{-1}$). This trend is due to the inside-out pattern of reionization imprinted by massive haloes, as discussed in Section 3.3.1. As their potential wells are large enough, photoheated gas is retained and remains as a hot, ionized, hydrostatically supported atmosphere in the CGM, where it may eventually cool and condense onto the central galaxy. The opposite is true for dwarf haloes, as these undergo outside-in reionization on average, hence potentially remain neutral until ionized by a distant source. Following their eventual photoevaporation, the only mechanism by which they are then able to accrete is via mergers and tidal stripping (as shown by the correlation between the yellow and dark blue points in the first and second rows of the third column, respectively).

Finally, in the bottom panels we show the virial ratio of haloes. At high-redshift, mergers are extremely common, therefore it is not always possible to discuss haloes as isolated systems (as is commonly the case when defining quantities based on the virial theorem). Haloes with excess kinetic energy (high virial ratio) have likely undergone strong interactions with a close neighbour, resulting in significant turbulence when baryons fall into the central regions (Davis et al., 2011). This may play a crucial role in the cooling and collapse of baryons in high-redshift haloes, as a higher ratio of velocity dispersion to the sound speed can lead to increased fragmentation (Clark et al., 2011; Greif et al., 2011). We find that haloes with excess kinetic energy typically have super-cosmic mean baryon fractions, however only account for approximately 10 per cent of the total population at $z \sim 6$, in good agreement with Davis et al. (2011). Over time, the virial ratio of low-mass haloes falls as they relax and reach virial equilibrium.

3.4 Artificial Neural Network

To explore the parameter space that influences the halo baryon fraction, we have developed an ANN, which we train using our fiducial simulation. The aim is to determine the minimum number of parameters required in order to accurately predict the simulation output, and more interestingly, make structured predictions based on variations in these parameters. Up to now, models fit the halo baryon fraction using a mass dependent fitting function (e.g. Henriques et al., 2015; Lacey et al., 2016). However, as we have

already shown in this work, the halo baryon fraction does not follow such a tight relation. Rather, a large amount of scatter is evident at all times, as one might expect, which can drastically alter the amount of available baryons for both galaxy and star formation. We demonstrate that by using ANNs and probing the impact of tidal forces, gas temperature, hydrogen ionization fraction, and virial ratio, one can accurately capture this scatter, construct a model that can reproduce these results, and drastically improve on existing methods.

Below, we briefly describe the architecture of our ANN and the choice of back-propagation algorithm used to train the model. We conclude this section by describing the adaptive-learning process we have implemented to improve the accuracy of our model in Section 3.4.3.

3.4.1 Architecture

Artificial neural networks are a branch of machine learning algorithms that mimic the natural neuron network in our brains through a series of mathematical weights, which are capable of learning any mathematical function to an accuracy proportional to the number of neurons in the network (Cybenko, 1989; Hornik et al., 1989). Many different ANN architectures exist, each suited to their own class of problems. At the simplest level, such a network consists of an input layer, one (or more) hidden layers, and a single output layer, each layer containing one or more neurons. Here, we develop an implicit, fully connected feed-forward network with a single hidden layer (shallow learning), as shown in figure 3.9, which is similar to that used in other works (see, e.g. Shimabukuro and Semelin, 2017).

The input layer consists of one neuron for each input parameter, which simply accepts the input value and outputs it to each neuron in the next layer (often referred to as a latch neuron). The linear combination of the inputs is then calculated for each neuron, s_i, in the hidden layer as:

$$s_i = \sum_{j=1}^{n} w_{ij}^{(1)} x_j + w_{i,b} b, \tag{3.3}$$

where neuron j receives the input data x_j and feeds it to the i-th neuron in the next layer with connection weight w_{ij}. The subscript b denotes a bias neuron, which has a fixed output value of unity in our case. The superscript refers to the layer number (where 1 denotes the hidden layer and 2 denotes the output layer, introduced below). The purpose of the bias neuron is to allow the network to vary its output along the input axis by modifying the weight $w_{i,b}$, such that the input values do not necessarily have to lie within

a given range, whilst the traditional connection weights w_{ij} control the steepness of the output.

Each hidden neuron is then activated by applying the transfer/activation function, which can take a linear or non-linear form, depending on the problem and desired behaviour of the network. We opt for a non-linear transfer function, which is then applied such that our neuron output becomes $h_i = f(s_i)$ and where the function $f(x)$ typically takes the form of a sigmoid function (as their derivatives are well known). The exact choice of transfer function again depends on the problem one desires to solve, as the training output must be bound by the output range of the neuron. For this work, we choose to use the hyperbolic tangent, however other choices may work equally as well. Each output neuron then sums the contributions from neurons in the hidden layer as above, such that:

$$z_i = \sum_{j=1}^{n} w_{ij}^{(2)} h_j + w_{i,b} b \tag{3.4}$$

and again the non-linear transfer function is applied to give the final output $y_i = f(z_i)$. While we have chosen to adopt both non-linear hidden and output layers, it should be noted that this is not strictly required, and one may prefer a non-linear transfer function for the hidden layer whilst simply taking the linear sum of weights for the output layer.

Finally, we may rewrite the above as a single equation, generalising the *forward-propagation* of inputs through the network as:

$$\begin{aligned} s_i^{(\ell+1)} &= \sum_{j=1}^{n} w_{ij}^{(\ell)} x_j^{(\ell)} + w_{i,b}^{(\ell)} b^{(\ell)}, \\ y_i^{(\ell+1)} &= f(s_i^{(\ell+1)}), \end{aligned} \tag{3.5}$$

where ℓ refers to layers 1 or 2 as above. We hereafter reserve the notation y_i (i.e without the super-script) for the output layer only.

3.4.2 Back Propagation

In this section, we discuss the backwards propagation algorithm used to train our ANN. For the network to quantify the accuracy of its predictions, we define the total cost function as:

$$E = \sum_{n=1}^{N} \left[\frac{1}{2} \sum_{i=1}^{m} (y_{i,n} - \hat{y}_{i,n})^2 \right] \tag{3.6}$$

where N is the total number of training samples, $y_{i,n}$ denotes the neuron output for this training pass, and $\hat{y}_{i,n}$ the target value. The network performs gradient descent at

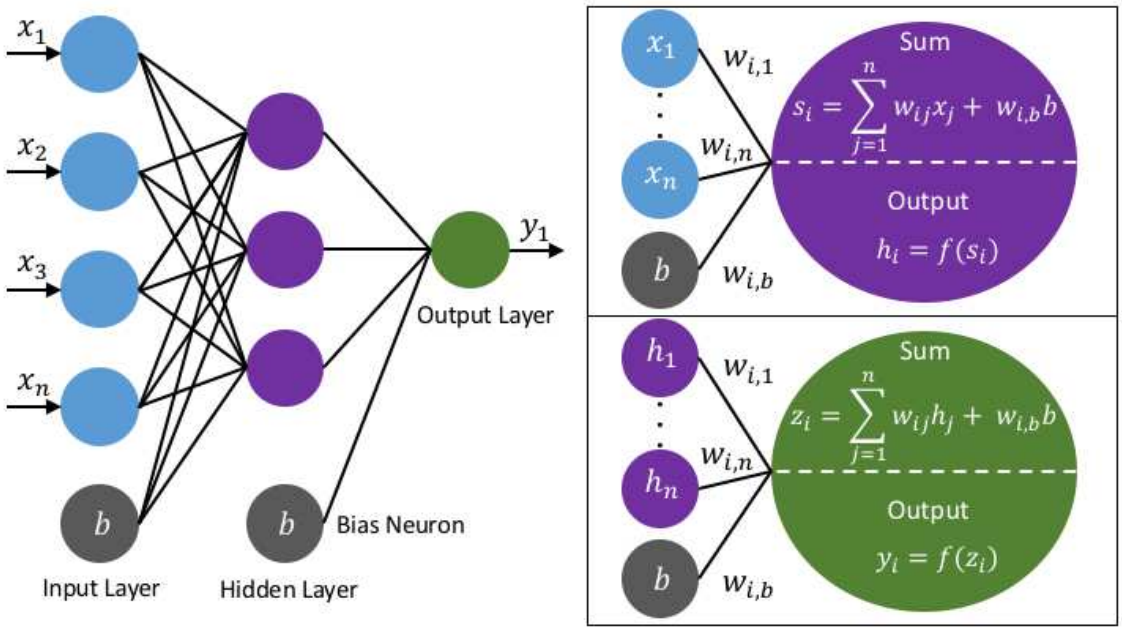

Figure 3.9: Illustration of the architecture of our ANN, showing the input, hidden and output layer neurons. Each neuron is connected to all neurons in the previous layer with its own mathematical weight, w, over which the contributions are summed and applied to the transfer function as shown in the boxes on the right. We opt for non-linear sigmoid transfer functions for both the hidden and output layers. The grey neurons in the input and hidden layer denote bias neurons, which have their own connection weight as above, however have a constant output value of unity. By optimizing the weights of these neurons, the network can vary the transfer function along the input axis, meaning we do not require our input variable to be bound by the transfer function (however some scaling does improve the learning rate of the model). The ANN performs gradient descent at each neuron connection to minimise the cost function (Eqn. 3.6) based on the target values supplied. We show three hidden neurons for illustrative purposes, however in practice we opt for a 4-45-1 topology (i.e 4 inputs, 45 hidden and 1 output).

each neuron connection to minimise this cost function by updating the internal system of weights, such that:

$$\begin{aligned} \Delta w_{ij}^{(\ell)} &= \eta \frac{\partial E}{\partial w_{ij}^{(\ell)}} + \xi \Delta \tilde{w}_{ij}^{(\ell)}; \\ \frac{\partial E}{\partial w_{ij}^{(\ell)}} &= \sum_{n=1}^{N} \frac{\partial E_n}{\partial w_{ij}^{(\ell)}}, \end{aligned} \tag{3.7}$$

where η and ξ are defined as the learning rate and momentum respectively and $\Delta \tilde{w}_{ij}^{(\ell)}$ is the change in the connection weight on the last training pass. These parameters are used to control how quickly the network learns and performs its gradient descent, which we allow to vary during the learning process and discuss in greater detail in the following section.

The derivative of the cost function with respect to the weights between the input and hidden layer can be calculated as:

$$\begin{aligned} \left(\frac{\partial E_n}{\partial w_{ij}^{(\ell)}} \right)^{(\ell)} &= \frac{\partial E_n}{\partial s_i^{(\ell)}} \frac{\partial E_i}{\partial w_{ij}^{(\ell)}} \\ &= \left[\frac{\partial E_n}{\partial y_i^{(\ell)}} \frac{\partial y_i^{(\ell)}}{\partial s_i^{(\ell)}} \right] x_j^{(\ell)} \\ &= \left[\delta^{(\ell)} f' \left(s_i^{(\ell)} \right) \right] x_j^{(\ell)}. \end{aligned} \tag{3.8}$$

Here, we have introduced the *error propagation* term $\delta^{(\ell)}$, where for the first layer:

$$\delta^{(1)} \equiv \sum_{n=1}^{m} \left(y_{i,n} - \hat{y}_{i,n} \right) w_{ij}^{(2)} \tag{3.9}$$

and m is defined as the number of neurons in the next layer. Note that in getting from equation (3.7) to equations (3.8) and (3.9) we have used the chain rule for the derivative, as E depends solely on the activated hidden neurons.

For the output layer, the error-propagation term is simply the difference between our network output and the target value multiplied by the connection weight:

$$\delta^{(2)} \equiv \left(y_{i,n} - \hat{y}_{i,n} \right) \tag{3.10}$$

A complete iteration to train the network therefore requires carrying out the following steps:

1. Initialise the network with random connection weights
2. Compute the outputs using equation (3.5) (Forward-propagation)

3. Compute the output gradients using equations (3.8) and (3.10)

4. Compute the hidden layer gradients using equations (3.8) and (3.9)

5. Update the set of internal weights using equation (3.7) (Back-propagation)

Once the network has been trained, the connection weights are exported and re-imported as required.

3.4.3 Adaptive Learning Rate

Generally, there is no standardised way of determining the appropriate learning rate and momentum of a neural network. One tends to use trial and error to determine the best choice of η and ξ, which varies based on the number of neurons and layers used. Typically, we want the initial learning rate to be fast, such that the network quickly determines a set of internal weights which roughly approximate the target function, without requiring a large number of training passes. However, this requirement could limit the final accuracy of the model, as large steps are taken in the gradient descent process thus making it difficult to converge on the global minimum. One solution to this problem is to allow the network to adaptively vary its own learning rate as the total root-mean-square (RMS) error falls below an acceptable threshold.

We opt for a simple adaptive learning implementation, whereby the parameters η and ξ are initially set to 0.15 and 0.25, and subsequently allowed to fall by a factor of two and four when the RMS error is measured to be consistently below some threshold. In figure 3.10 we show the resulting performance of the network with/without adaptive learning, where the error bars show the standard deviation in the RMS for each bin. The input data consist of a few thousand haloes, which the network makes several passes over in a random order such that the total number of iterations exceeds 10^6. As expected, the initial performance is identical in both cases, however the adaptive learning rate leads to a subtle reduction in the mean RMS after roughly 10^4 training passes. While the effect appears small, the adaptive learning rate allows the network to achieve a higher accuracy with sparse data sets and thus greater accuracy when capturing rare outliers in the data

3.4.4 Constraining M_c with Artificial Neural Networks

In line with recent studies (Noh and McQuinn, 2014; Okamoto et al., 2008; Hoeft et al., 2006), we quantify the scale at which the baryon fraction transitions from poor to rich

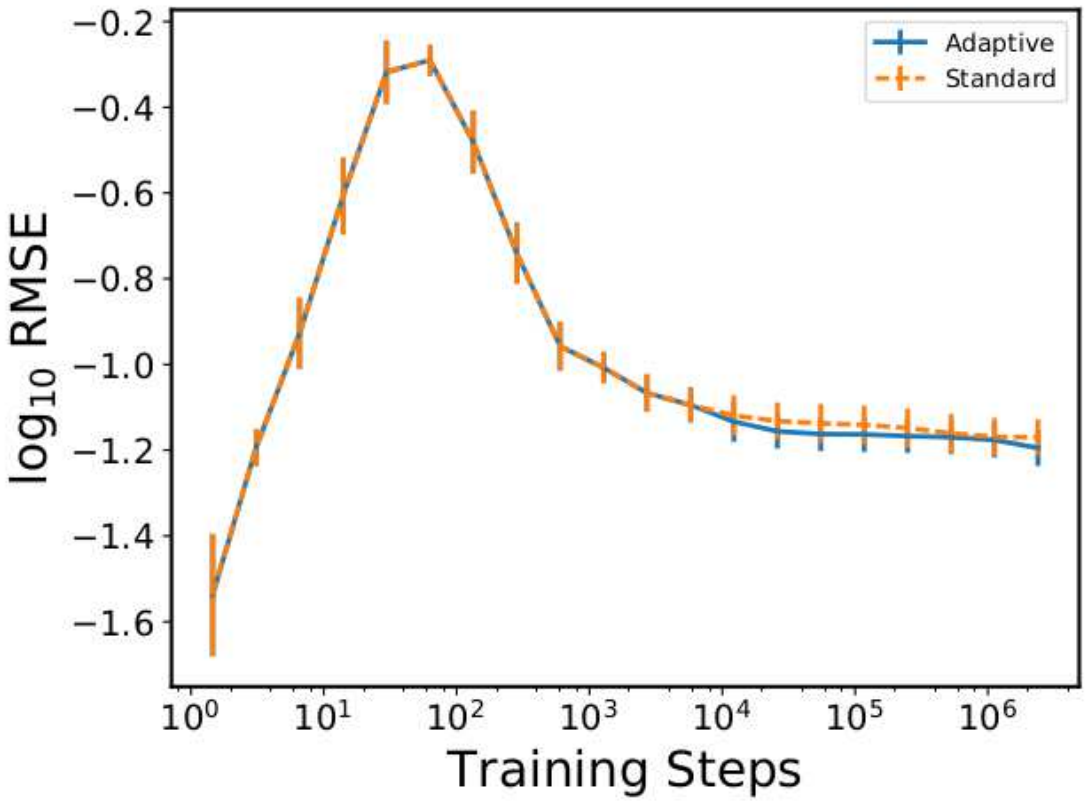

Figure 3.10: RMS error of the ANN, for the adaptive (varying; blue) and standard (fixed; orange) learning rates. Both models produce identical errors for the first 10^3 iterations, after which the adaptive method begins to reduce the net learning rate η and momentum ξ, allowing the gradient descent to get closer to a global minimum. While the difference appears small, this mostly corresponds to improved accuracy at capturing rare outliers in the data.

using the Gnedin (2000b) fitting formula:

$$f_{\rm b}(M,z) = \langle f_{\rm b} \rangle \left\{ 1 + \left(2^{\alpha/3} - 1 \right) \left[\frac{M}{M_{\rm c}(z)} \right]^{-\alpha} \right\}^{-3/\alpha}, \tag{3.11}$$

where $\langle f_{\rm b} \rangle$ denotes the cosmic mean baryon fraction, M is the halo mass, $M_{\rm c}(z)$ is the redshift dependent characteristic mass scale defined earlier, and α is an exponent which controls the steepness of the transition between the two extremes. The characteristic mass scale represents a halo whose baryon fraction is half the cosmic mean value, as determined from hydrodynamical simulations and necessary to fit the baryon fraction in semi-analytical models of galaxy formation (such as the Munich and Durham semi-analytical models; i.e. L-GALAXIES and GALFORM, respectively; Henriques et al. 2015; Lacey et al. 2016). While the exponent need not be constant, a value $\alpha = 2$ is found to give good agreement with our results.

It was recently found by Aubert et al. (2015) that $f_{\rm b}$ is independent of reionization history at fixed *volume-weighted* ionized fraction. This is significant, as we have departed from thinking about how $f_{\rm b}$ varies in time (redshift), to instead considering how it varies based on the physical evolution of haloes. This paradigm shift is our primary motivation in developing an ANN to understand, and predict this quantity based solely on time-independent physical quantities.

We train the model with our fiducial simulation only, using the four quantities shown in figure 3.7 as our input parameters. The training data consists of haloes from eight snapshots between $z \sim 12$ and $z \sim 6$, which we write out randomly over several thousand passes such that the total number of training samples exceeds 10^6. Once trained, we export the internal system of connection weights, which is all that is required to calibrate the model for future use. We use this model to predict $f_{\rm b}$ based on our simulated halo catalogues, to test the accuracy of the network, and present our results in figure 3.11 for our fiducial simulation (left column) and ANN (right column) at $z = 6.15$.

Our model reproduces the simulated data well for all quantities, matching the trends for each quantity except for the rarest objects. These represent only a few percent of the total halo population, and require additional physics in our training data to fully capture. While the scatter is reduced in our predictions, as shown by the black dashed fit and error bars, it is a remarkable improvement over existing constraints which are mass and redshift dependent only. The network captures many interesting properties from our simulated haloes, such as the correlation between high tidal forces and large neutral fractions at low-masses and the presence of high tidal forces above and below the mean, with relaxed haloes following the best fit curve. The overall accuracy could be further improved by

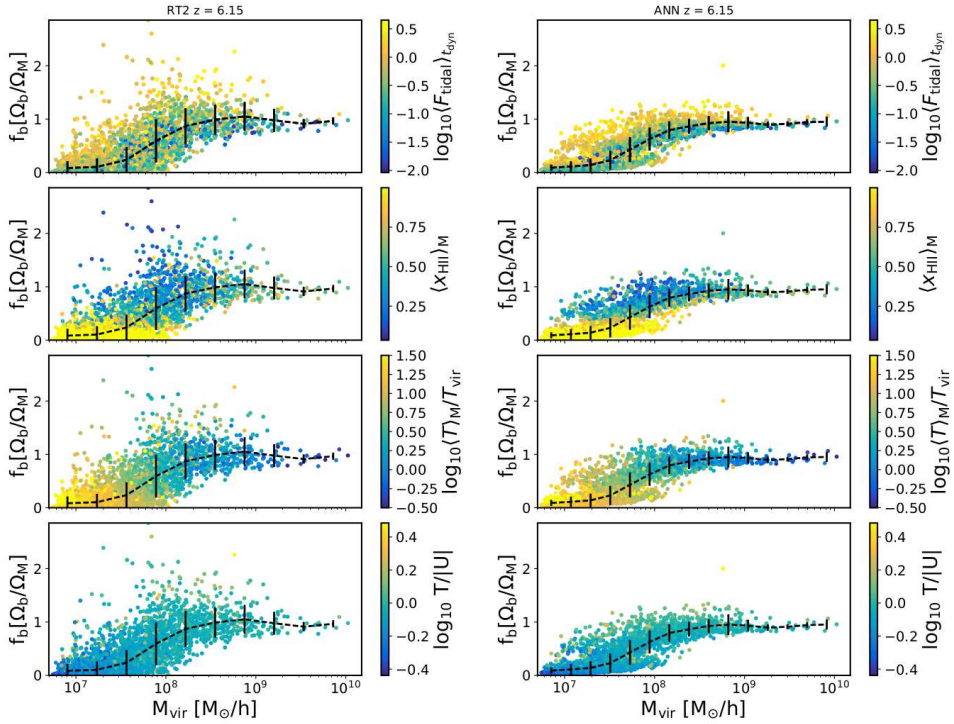

Figure 3.11: Same as figure 3.7, but for our fiducial model (left) and ANN predictions (right) at $z = 6.15$. We use the existing halo catalogue data, as it is presented in the left-hand column, for our network predictions. The ANN is trained using only our fiducial simulation, and is then tested against the remaining RT1.5 and RT5 models with M_c predictions shown in figure 3.12 (shaded is simulated, error bar points are our ANN predictions). The model shows extremely good agreement for all quantities shown, however struggles to match the most extreme cases (only a few percent of haloes) as we train over all redshifts where the degree of scatter varies over time, therefore the network smooths over the rarest objects.

including more input parameters, to increase the number of degrees of freedom in the model, however we restrict our work to a limited sample as we aim to improve usability in SAMs of galaxy formation; which may not be able to quantify a large range of inputs accurately.

To probe accuracy over a range of redshifts, we compute $M_{\rm c}(z)$ by non-linear least squares fitting to Eqn. 3.11 for a range of snapshots between $12 \leq z \leq 6$. The exponent, α, is left as a free parameter, however a value of two is found to give good agreement at all times. This is done for all simulations listed in Table 3.1 (shaded) and the corresponding ANN predictions at the same redshifts (points), as shown in the top left panel of figure 3.12. Our fiducial simulation falls significantly below the predictions of Okamoto et al. (2008) (black dot-dashed line) shortly after their instantaneous reionization at $z_{\rm reion} = 9$, and above at earlier times. This is due to the inhomogeneous nature of reionization in our simulations, as H II regions begin to grow around the first stars which form as early as $z \sim 18$ (see figure 3.5). While this leads to an increased suppression of the halo baryon fraction prior to the EoR, as $M_{\rm c}(z)$ is significantly larger, we find that haloes are not as strongly suppressed as previously thought following its completion. In fact, only our early reionization simulation, in which haloes are suppressed much earlier in their lifetimes, gives good agreement by $z \sim 6$. We also show the predictions from Hoeft et al. (2006), who use a similar homogeneous UV background model as discussed above, but trigger instantaneous reionization at $z_{\rm reion} = 6$; hence their constraints initially lag behind before converging at the present day ($z = 0$).

The ANN, trained on our fiducial haloes only, shows extremely good agreement at all times, with the uncertainties becoming smaller with redshift in line with our simulations. As $M_{\rm c}$ sets the mass scale where $f_{\rm b}$ transitions from a few percent to the cosmic mean, this shows that our model can accurately predict the shape of this relation independent of the underlying reionization history, with the introduction of physical scatter in our $f_{\rm b}$ predictions. This represents a significant improvement over existing constraints on $M_{\rm c}(z)$, which are model, mass, and time dependent. To demonstrate that our findings are independent of redshift, we show the same quantity but as a function of *mass-weighted* hydrogen ionized fraction in the top right panel of figure 3.12. We find good agreement with Aubert et al. (2015), however note that they compare the halo baryon fraction as a function of *volume-weighted* ionized fraction. When investigating the internal properties of haloes, we are primarily concerned with the mass structure over the total volume, hence we opt for the mass-weighted quantity. This demonstrates that the Jeans filtering

of haloes during the EoR is a *local* effect, and not globally time dependent. Therefore, redshift dependent constraints on $M_{\rm c}$ are not well motivated physically, and miss much of the scatter in the galaxy population during this time.

In the bottom two panels of figure 3.12, we show the same as above but with a tidal cut-off, whereby we exclude all haloes where $\langle F_{\rm tidal} \rangle_{t_{\rm dyn}}$ exceeds the value of the highest peak in the PDF at each redshift (figure 3.8). Okamoto et al. (2008) were similarly motived to remove haloes which had undergone recent merger activity to reduce the super-cosmic mean scatter in their simulations. As demonstrated in figure 3.7, the net effect of tidal stripping is to raise the baryon fraction above the halo average over time (top row, left to right: high density of yellow points above the dashed mean). Therefore, we find that excluding such objects leads to a significant rise in $M_{\rm c}$ for all models. To test the sensitivity of our ANN to such a scenario, we create a mock training sample whereby we re-sample the tidal forces exerted on all haloes from the PDFs shown in figure 3.8, but below the large peak as done above. The model correctly predicts a rise in $M_{\rm c}$ for all simulations, with minor tension shown for RT1.5. This tension is likely due to the random re-sampling of the PDF, however the model still accurately predicts the distribution of all other quantities (see Appendix 3.B, figure 3.15). In the bottom right panel, we find that all models are still consistent with each other as a function of mass-weighted ionized fraction, suggesting that photoionization/heating are the dominant processes in controlling $f_{\rm b}$.

3.5 Discussion

The Jeans-mass filtering of low-mass galactic haloes embedded within H II regions has long been proposed as an efficient mechanism to regulate star formation, by preventing the accretion of baryons from the photoheated IGM (e.g Shapiro et al., 1994; Quinn et al., 1996; Gnedin, 2000b). Haloes that achieve a high stellar emissivity will reach a much higher ionized fraction earlier than those with low star-forming efficiencies, which in turn will lead to them undergoing Jeans-mass filtering earlier. This effect compensates for the difference in emissivity of haloes in the suppressible regime, leading to a 'self-regulating' reionization (Iliev et al., 2007a), and could help explain the discrepancy between the observed number density of Milky Way dwarf satellites and those predicted by N-body simulations that assume a ΛCDM cosmology. However, previous studies on this subject have failed to intricately resolve the internal structure of galactic haloes during reionization, potentially overestimating the efficiency of radiative feedback on star formation.

Semi-analytical models predict that the modulation of star formation due to Jeans-

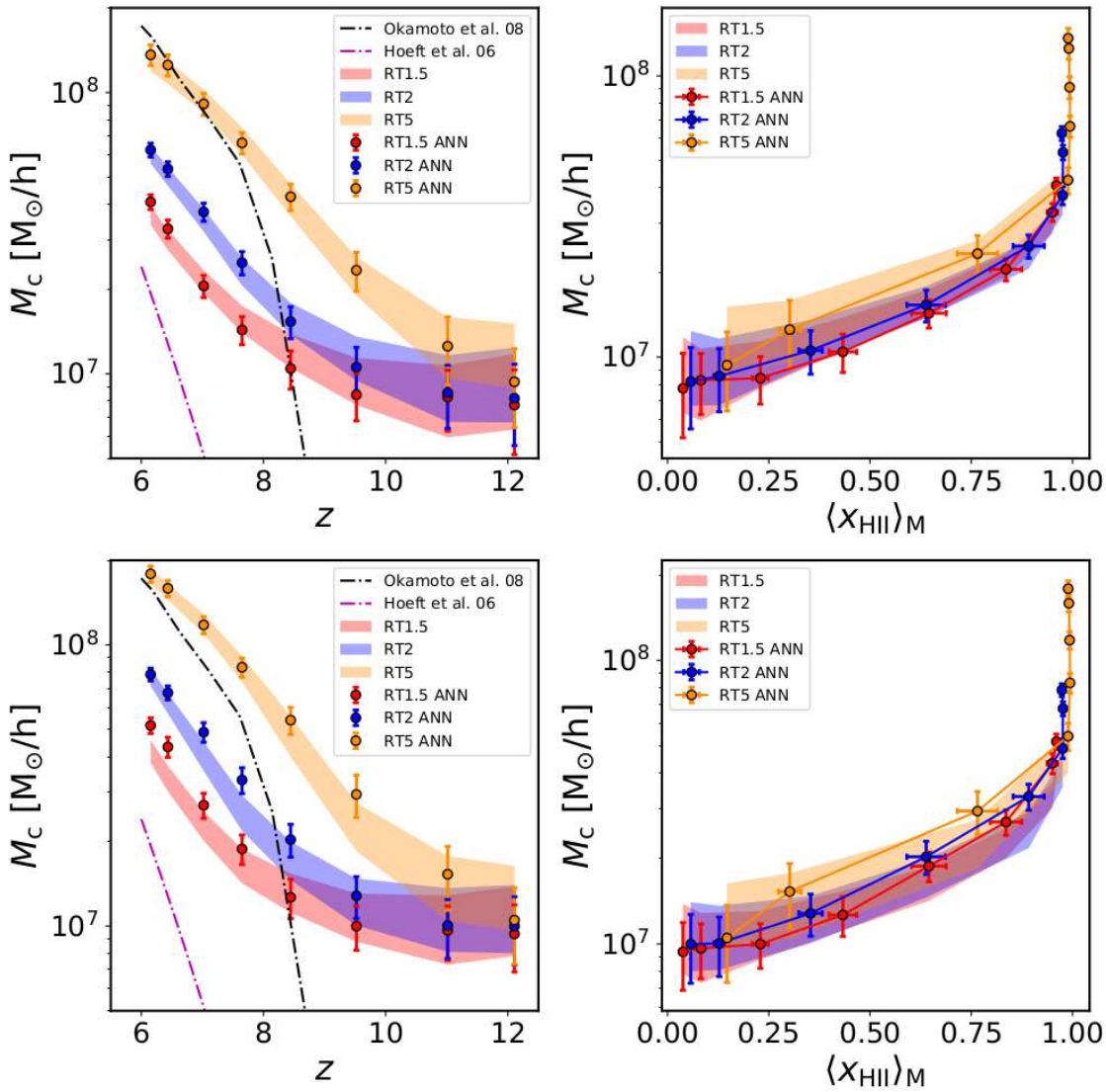

Figure 3.12: *Top left panel:* Evolution of the characteristic mass scale, M_c, as a function of redshift for all RTX models listed in Table 3.1 (shaded), and our ANN predictions (points). The error bars correspond to the 5σ confidence level. The dot-dashed lines show the fits from Okamoto et al. (2008) (black) and Hoeft et al. (2006) (magenta), who assume a quasi-homogeneous UV background using the models of Haardt and Madau (2001) and Haardt and Madau (1996) respectively. *Top right panel:* Evolution of M_c as a function of mass-weighted hydrogen ionization fraction, showing that all models are consistent at fixed points in their reionization histories. Note, we do not include the results of Okamoto et al. (2008) and Hoeft et al. (2006) as their models transition from a neutral to fully-ionized Universe instantaneously (step function). *Bottom panels:* Same as above, but with a cut-off in the tidal force below the position of the highest peak in the redshift dependent PDF (see figure 3.8) applied to the simulated data. While the ANN predictions for RT1.5 in our tidal cut-off model (bottom left panel) are slightly high, overall the network matches the simulated data extremely well.

mass filtering at the reionization redshift should be observable in both the cosmic star formation rate of low-mass galaxies and the faint end of the UV luminosity function (e.g Barkana and Loeb, 2000). Recent RHD simulations of the EoR have also drawn this conclusion, showing that the comoving star formation rate of the lowest mass haloes, below masses where atomic cooling is possible, are significantly suppressed and in some cases can even be quenched entirely (e.g Ocvirk et al., 2016). Large scale N-body + RT simulations, such as those by Iliev et al. (2007a) and more recently Dixon et al. (2016), which probe the global geometry and statistics of reionization, incorporate these assumptions into their models as sub-grid physics in order to determine which sources drive reionization.

High-resolution simulations of isolated disk galaxies, such as those carried out by Rosdahl et al. (2015), have shown that radiative feedback primarily acts to prevent the collapse of gas and thus the formation of star-forming clumps, rather than destroy those that have already formed. Rosdahl et al. (2015) conclude that the reduction in star formation due to radiative feedback ranges from a factor of 4 in low-mass galaxies to roughly 0.1 for the most massive, similar to that of traditional thermal dump SNe feedback models. This reduction is much weaker than recently predicted by large-scale RHD simulations of the EoR and assumed in the sub-grid models of N-body + RT simulations, such as those discussed above.

One of the primary aims of this paper was to determine the statistical effect of radiative feedback on galaxies in a relatively small cosmological volume. In Section 3.3.1, we carried out a statistical analysis of the distribution of gas densities and temperatures within haloes and concluded that while radiative feedback can heat the gas above our temperature threshold for star formation, it did not have a significant impact on galaxies, as gas at typical ISM densities can already cool efficiently. The CDF of the neutral gas of haloes exposed to ionizing radiation in our simulations show a large reduction in the fraction of dense gas at all masses with respect to the SNe-only case, with low masses most strongly effected. These lower on average densities lead to a reduction in the star formation rate and stellar abundance of haloes; however, low-mass haloes are not completely quenched. As such, we conclude that radiative feedback is much weaker than previously thought, in good agreement with studies of isolated disks (e.g Rosdahl et al., 2015).

To understand why the lowest-mass galaxies are able to continue forming stars in our simulations, we investigated how various physical effects impacted the baryon fraction of haloes, $f_{\rm b}$. These included tidal forces, photoionization, heating, and the degree to which haloes are virialised at high redshifts. The Jeans-mass filtering and subsequent

photoevaporation of the lowest-mass haloes leads to a reduction in $f_{\rm b}$ to a few percent of the total halo mass, while intermediate to high-mass haloes are able to retain a significant fraction of their baryons. This effect becomes more exaggerated towards to end of reionization, as photoevaporation occurs on time-scales of a few hundred million years (Shapiro et al., 2004; Iliev et al., 2005b). The highest-mass haloes remain at the cosmic mean baryon fraction, despite radiation preventing collapse and keeping gas at low densities, as previously demonstrated by the CDF of neutral hydrogen. Their large potential wells can retain photoheated gas, remaining as an ionized, hot, hydrostatically supported atmosphere. Tidal stripping leads to a large scatter in $f_{\rm b}$ above the median, allowing the lowest-mass haloes to accumulate significant fractions of neutral gas. If this gas can cool and collapse to high densities, it may become self-shielding from ionizing UV radiation and thus allow prolonged star formation. The ability for the gas to become self-shielding requires extremely high-resolution simulations, which resolve gas at extremely high overdensities ($\rho/\langle\rho\rangle \gg 10^5$). At these high densities, self-shielding could also enable the formation of cold accretion streams, which allow massive galaxies to continue to infall baryons throughout reionization (see Appendix 3.A and Rosdahl and Blaizot 2012). Lower resolution simulations which do not resolve these scales are therefore prone to over-suppressing galaxies, as I-fronts can very quickly sweep through the gas and indefinitely prevent collapse until the potential well grows sufficiently large through merging.

The same is found by comparing numerical constraints on the characteristic mass scale, $M_{\rm c}$, with those of Okamoto et al. (2008) whose numerical fits are commonly used to constrain $f_{\rm b}$ in SAMs of galaxy formation. In their models, homogeneous reionization is triggered instantaneously at $z_{\rm reion} = 9$, using the UV background model of Haardt and Madau (2001). This assumption leads to a steep increase in $M_{\rm c}$, as haloes transition from unsuppressed to suppressed regime much faster than in our self-consistent models, where their suppression is dependent on their *local* reionization history. The suppression continues to rise above our fiducial model due to a number of factors. First, contrary to our models, they neglect metal-line cooling, which denies the gas an efficient cooling mechanism at high densities. Secondly, due to their fixed number of smoothed particle hydrodynamics (SPH) particles, the smoothing length is only $\epsilon = 0.25$ kpch^{-1}. This choice results in a maximum baryon overdensity of $\rho/\langle\rho\rangle \sim 10^5$ in their simulations, while we achieve $\rho/\langle\rho\rangle \sim 10^8$, which could reduce the ability of gas to self-shield efficiently. Finally, haloes which have undergone recent merger activity are removed from their results prior to fitting $M_{\rm c}$. As we have shown, strong tidal interactions leads to a large scatter, typically

above the median, in $f_{\rm b}$. Removing these haloes from our sample leads to an increase in $M_{\rm c}$, partially easing the tension with our fiducial model. In contrast, our fiducial results are several orders of magnitude above the predictions of Hoeft et al. (2006), despite using the same SPH code GADGET2 (Springel, 2005) and similar homogeneous UV background model of Haardt and Madau (1996) compared to that of Okamoto et al. (2008). In their model, reionization is not triggered until $z_{\rm reion} = 6$, hence their values initially lag behind those discussed above. Furthermore, they restrict their study to void galaxies, which are not representative of the wider galaxy population.

To improve on their models, we developed an ANN to predict the halo baryon fraction that is independent of redshift, mass, and crucially, reionization history. This model can account for the impact of mergers and tidal stripping, as we demonstrated by creating a mock catalogue of haloes with low tidal forces attributed to them. Both our simulations and ANN model predict a rise in $M_{\rm c}$ in such a scenario; however, there is mild tension between them due to the manner in which we randomly sample tidal forces to construct out training data. While the ANN reproduces much of the scatter in $f_{\rm b}$, marking a significant advancement over the current methods of SAMs that use only a mass- and redshift-dependent best-fit curve, it misses a small percentage of haloes with the highest baryon fractions. These objects require additional physics to constrain them within our model, which in turn would introduce additional degrees of freedom. One caveat of our model is that it requires the mass-weighted ionized fraction and gas temperature to be calculated for each halo, which may not be possible in some simpler SAMs. Therefore, we restrict the number of inputs (degrees of freedom) to four, to reduce the number of known quantities required to implement our model.

While our simulations are consistent with observational constraints of the volume-weighted hydrogen neutral fraction, photoionization rate, and Thompson scattering optical depth, there is the caveat that relatively small volumes exhibit a significant scatter in their EoR when examined as part of large-scale simulations (see, for example Ciardi et al., 2003a; Iliev et al., 2006b; Mellema et al., 2006b; Zahn et al., 2007; McQuinn et al., 2007; Iliev et al., 2007a; Alvarez and Abel, 2007; Mesinger and Furlanetto, 2007; Geil and Wyithe, 2008; Choudhury et al., 2009). While RAMSES-RT employs a moment based method of RT, which scales independently of the number of sources and hence is not limited by volume, the code does not scale particularly well at our resolution of $\Delta x = 15.25$ pc h^{-1}. As such, the limiting factor in scaling our simulations up to larger volumes lies solely with access to large computational resources.

Due to computational constraints, we follow the methodology of Gnedin and Abel (2001) in reducing the speed of light to $f_c = 1/100$. This assumption allows us to relax the Courant-condition for the radiation fluid, thus reducing the total number of radiation sub-cycles and, hence, the total run-time (see also discussion in Aubert and Teyssier 2008). A framework for setting the reduced speed of light is set out in Rosdahl et al. (2013), following the study of relativistic I-fronts in Shapiro et al. (2006). A reduced speed of light increases the effective crossing of I-fronts, hence they will lag behind the full speed of light solution after a single light-crossing time (across the Strömgren radius, in units of the recombination time). This effect can cause a retardation in the speed of I-fronts across cosmological volumes; therefore, we artificially boost the stellar escape fraction, $f^*_{\rm esc}$, in order to calibrate our reionization histories as shown in figure 3.2. While our minimum value is above unity, i.e. $f^*_{\rm esc} = 1.5$ for our RT1.5 simulation, we note that stellar synthesis uncertainties could give rise to much higher stellar luminosities than those predicted by Bruzual and Charlot (2003) and used in our simulations (see the BPASS model, for example; Eldridge et al. 2008). As we are primarily concerned with the properties of individual haloes in our simulations, and not the global statistics of reionization, the reduced speed of light approximation is valid as the effective crossing time of a halo is negligible when compared with the Universe.

3.6 Conclusions

We have presented the highest spatially resolved, fully self-consistent simulations of the halo baryon fraction during the EoR, achieving a formal resolution of 15.25 pch^{-1} comoving using the RAMSES-RT code (Rosdahl et al., 2013). To disentangle the effects of tidal stripping, photoionization, and heating, we developed an Artificial Neural Network, which we trained using our fiducial simulation and then applied to three distinctly varying reionization histories. Our network is able to accurately reproduce the halo baryon fraction in each model, matching the measured value of the characteristic mass scale, $M_{\rm c}$ in each case (figure 3.12).

Our findings are in disagreement with those of Okamoto et al. (2008) and Hoeft et al. (2006), both of whom adopt the (instantaneous) homogeneous UV backgrounds of Haardt and Madau (2001) and Haardt and Madau (1996), respectively. This discrepancy is partially due to stronger self-shielding in our models (as we reach much higher spatial resolutions and hence baryon overdensities) and an additional cooling mechanism available (i.e. efficient metal-line cooling at high densities). Our most extreme RT5 model, which is in

tension with observations of the neutral fraction and Thompson scattering optical depth constraints from Planck (Planck Collaboration et al., 2016a), does match the Okamoto et al. (2008) constraints by $z \sim 6$ and overshoots this value when haloes that have undergone tidal interactions are ignored.

Radiative feedback can suppress star formation in haloes, however not as strongly as previously suggested by Ocvirk et al. (2016). While the star formation rate in dwarf haloes (below the atomic cooling limit) is reduced, it is not quenched entirely. Furthermore, while the mean stellar abundance of haloes is reduced in our RT models, only our early reionization model (RT5) is inconsistent with the reference SNe-only case within 1σ. The degree of suppression in SFR is strongly correlated with the reduction in halo baryon fraction, as radiation primarily acts to prevent the collapse of star forming gas, rather than heat it past the temperature criterion for star formation. We find that modulation of star formation is primarily due to Jeans-mass filtering, preventing the collapse of baryons from the IGM into haloes and thus reducing their baryon fraction

The primary aim of this paper was to develop a model that can predict and capture the scatter in the baryon fraction of haloes in the suppressible regime, whose star formation rates are modulated during reionization. Existing models rely on fitting the average mass scale and slope that determines the transition from baryon poor to rich, providing a best-fit curve that misses much of the scatter in this quantity. We developed and trained an artificial neural network (ANN) to reproduce the halo baryon fraction based on four input parameters: the strongest tidal force exerted from a neighbour (averaged over a dynamical time), the (mass-weighted) hydrogen ionization fraction, the gas temperature (in units of the halo virial temperature), and the virial ratio (ratio of kinetic to potential energies). With these limited degrees of freedom, we demonstrated that our ANN could reproduce the characteristic mass scale, $M_{\rm c}$, independently of redshift, mass, and reionization history. Furthermore, it captures much of the scatter in the baryon fraction, leading to a more diverse galactic halo population.

Two of our input parameters (tidal force and virial ratio) can be computed purely from N-body simulations, and both are already provided by the ROCKSTAR halo finder (Behroozi et al., 2013b). The remaining two quantities, the hydrogen ionization fraction and gas temperature can be approximated by semi-analytical models, see e.g. Srisawat (2016). Therefore, our ANN can be trivially embedded within semi-analytical models of galaxy formation, leading to a more self-consistent calculation of the baryon fraction over current methods.

Acknowledgements

The authors contributed in the following way to this paper. DS undertook the calibration and completion of all simulations; the development of all analysis codes, including the artificial neural network; the production of all figures and the first draft of the paper. II supervised the project and helped with the interpretation of results and feedback on the first draft. KD provided guidance throughout and significant feedback on the first draft.

DS would like to thank Joakim Rosdahl for valuable technical feedback, Scott J. Clay for advice and feedback on the first draft, Peter A. Thomas, Paul R. Shapiro and Benoît Fournier for useful discussions and feedback.

This work used the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The DiRAC system is funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grant ST/H008519/1, STFC DiRAC Operations grant ST/K003267/1, and Durham University. DiRAC is part of the National E-Infrastructure. We acknowledge PRACE for awarding us computational time under project PRACE4LOFAR grant 2014102339 and "Multi-scale simulations of Cosmic Reionization" grants 2014102281, 2015122822, and 2016153528 to resource Curie based in France at CEA and to resource Marenostrum based in Spain at BSC. This work was supported by the Science and Technology Facilities Council [grant numbers ST/F002858/1 and ST/I000976/1] and the Southeast Physics Network (SEPNet). Some of the analysis was done on the Apollo cluster at The University of Sussex.

We made extensive use of the PYMSES (Guillet et al., 2013) and PYNBODY (Pontzen et al., 2013) python modules when developing our own analysis pipeline.

Appendix

3.A Cold Accretion Streams

Cold (neutral) accretion streams can form when gas reaches sufficiently high densities that it is able to self-shield from ionizing UV radiation (where recombinations balance out ionizations and equilibrium is achieved). Resolving these phenomena in fully coupled radiation hydrodynamical simulations is expensive, as one must self-consistently account for the hydrodynamical response to feedback whilst capturing these small scales in both galaxies and filaments. While some studies focus on the accurate treatment of the RT at the expense of capturing feedback (Nakamoto et al., 2001; Razoumov and Sommer-Larsen, 2006; Ciardi et al., 2003b; McQuinn et al., 2007; Finlator et al., 2009; Aubert and Teyssier, 2010), others ignore the dynamics of the gas altogether (Iliev et al., 2006b), using a semi-analytical approach to calibrate the sub-grid response (Iliev et al., 2007a).

More recently, improvements in algorithms and the steady increase in available computing power has lead to the first fully self-consistent radiation-hydrodynamical simulations of galaxy formation to resolve this process (Rosdahl and Blaizot, 2012; Pawlik et al., 2015; Ocvirk et al., 2016). To demonstrate the existence of self-shielding in our models, we focus on a single halo with mass $\mathrm{M_{vir}} \sim 3 \times 10^9\ M_\odot h^{-1}$ and virial radius $R_{\mathrm{vir}} \sim 8$ kpc at $z \sim 6.4$ in our fiducial simulation. We show projections of gas column density (FFT convolved map), minimum temperature, and minimum hydrogen ionization fraction along rays cast through a sub-volume of 40 kpc^3 surrounding the halo (figure 3.13, top to bottom respectively). The steel-blue circle annotates the boundary of the virial sphere; while in the top panel inset, we show a zoom-in of the central galaxy, re-sampled at the full resolution. In the centre and bottom panel, we see the accretion streams have a (relative) cold neutral core, which becomes disrupted as it passes the virial sphere (shock heated to the virial temperature). These streams allow the transport of cold, neutral gas from the IGM to the halo, where it may cool and form stars (provided it continues to self-shield once reaching the halo centre).

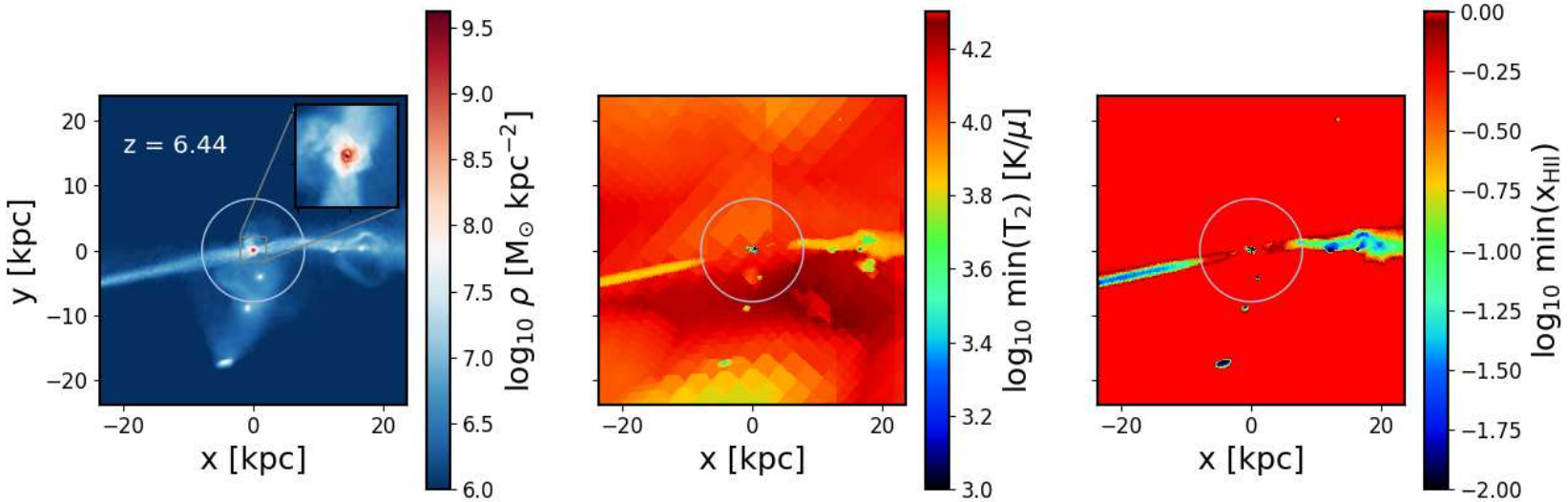

Figure 3.13: Projections of a 40 kpc^3 volume surrounding one of our RT2 haloes at $z \sim 6.4$ with mass $\mathrm{M_{vir}} \sim 3 \times 10^9\ M_\odot h^{-1}$ and virial radius $R_{\mathrm{vir}} \sim 8$ kpc. *Top panel:* Gas column density, showing the central galaxy and accretion streams. We zoom-in on the galaxy and re-project it at our full resolution. The steel-blue circle denotes the boundary of the virial sphere. *Middle panel:* Minimum temperature in units of K/μ, calculated by ray-tracing the volume with 512^2 rays. As gas passes the virial radius, it is shock heated to the virial temperature of the halo. *Bottom panel:* Same as above, but for hydrogen ionization fraction, showing dense neutral cores within the accretion streams. This gas is sufficiently dense to self-shield from the background UV radiation.

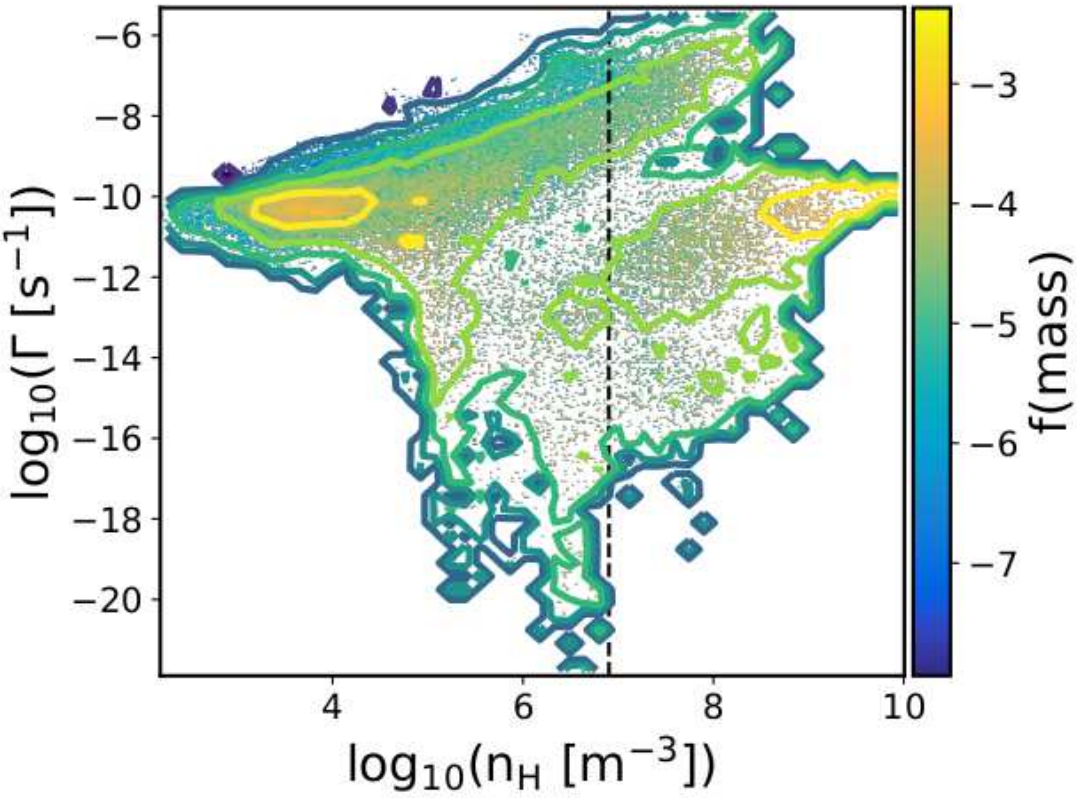

Figure 3.14: Mass-weighted photoionization rate-number density 2D histogram of the same volume shown in figure 3.13. The black dashed line denotes the density criterion for star formation, n_*. As the density increases, the photoionization rate drops substantially, as gas is able to self-shield from ionizing UV and balance recombinations and ionizations. The photoionization rate increases at densities $n_{\rm H} > n_*$, as stars forming in the high density peaks, preferentially ionizing these regions.

The effect of self-shielding is made clearer by visualizing the distribution of cells in photoionization-density space, as shown in figure 3.14. The black dashed line denotes the density criterion for star formation, n_*, while the colour-bar represents the fraction by mass of cells that occupy this state. As the gas density increases, the photoionization rate decreases substantially due to self-shielding. It then increases at densities $n_{\rm H} > n_*$ as these regions under photoionization from the inside-out in concordance with our reionization geometry (see Figs. 3.1 and 3.2). This effect is also seen in Ocvirk et al. (2016), however to a lesser degree due to their limited numerical resolution (and hence maximum density achieved).

3.B Low Tidal Force Model

In this appendix, we show the results of our low-tidal force model (as described in the text). As shown in figure 3.11, strong tidal interactions act to increase the scatter in $f_{\rm b}$, especially at the low-mass end. To test the predictive power of our ANN, we re-sample the tidal forces of all haloes using the PDFs shown in figure 3.8, whereby we place a cut-off at the position of the mean for the largest peak. For our simulated haloes, approximately half of the total populations exhibit tidal forces below this threshold, however our aim here is to develop a simple toy model to test if our ANN provides sensible predictions (i.e. has it properly understood the relationship between tidal force and baryon fraction).

In figure 3.15, we show the original halo catalogue results for our fiducial simulation (left column, as shown previously) and our ANN predictions for our test model (right column). Note, we re-scale the colour-bar for the top row for visualization purposes only. Our test model behaves as expected, leading to a significant reduction in the $\rm M_{vir}$-$f_{\rm b}$ relation and also a rise in the characteristic mass (figure 3.12, bottom panels). As discussed in the text, Okamoto et al. (2008) use a similar approach to reduce scatter in their constraints, whereby they remove all haloes that have undergone recent merger activity. Exclusion of such objects leads to an increases in the numerical values for $M_{\rm c}$, as we have shown.

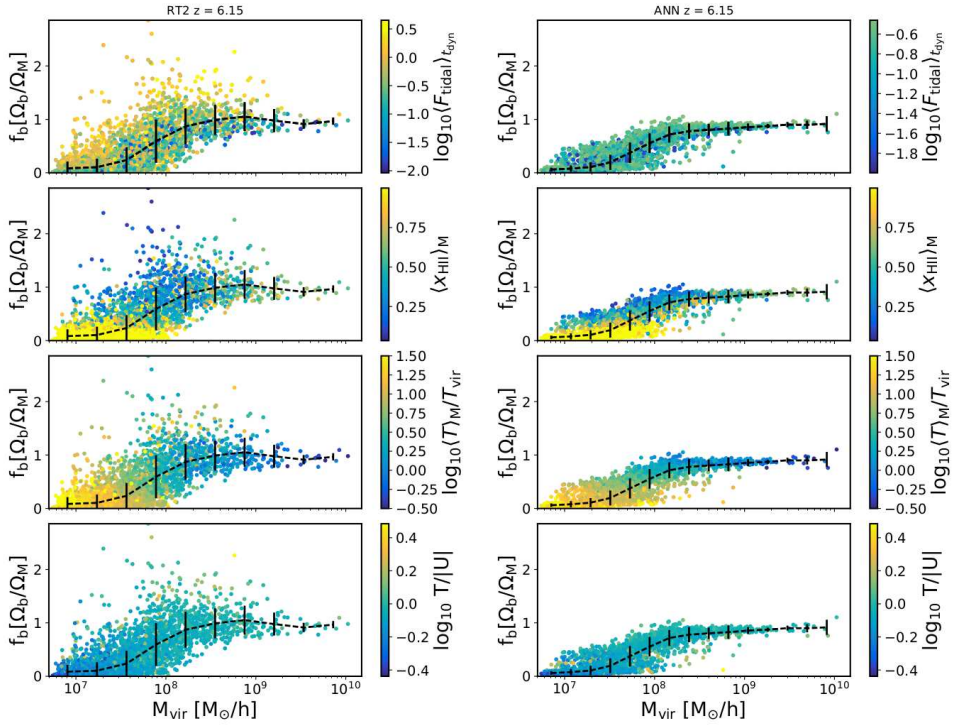

Figure 3.15: Same as figure 3.11, but for our low-tidal force model. We re-sample the tidal force of each halo using the PDFs shown in figure 3.8 and assign a new value below the position of the largest peak. Our ANN is then used to predict the halo baryon fraction for such a population, as shown in the right-hand column (note: the left-hand column is unchanged from figure 3.11, for reference). A reduction in the tidal force exerted on all haloes results in a tightening up of the scatter in f_b and an increase in our non-linear least squares fitting of M_c (see figure 3.12, bottom panels).

Chapter 4

Impact of Early Binary Populations on Radiatvie Feedback and the Escape Fraction of Ionizing Photons

In this chapter, we investigate the impact of stellar evolution uncertainties, namely whether or not binaries are accounted for, on the escape fraction of ionizing photons. The escape fraction is a crucial parameter in any model of reionization, however remains poorly defined and understood. Recent studies have suggested that an enhancement in the escape fraction of ionizing photons due to binary stellar populations can account for the full photon budget required to ionize the IGM with galaxies as the only sources. As these studies rely on semi-numerical techniques, we investigate whether binaries do indeed boost the escape fraction using fully-coupled radiation hydrodynamical simulations of the EoR. We find that while the time-integrated escape fraction is enhanced, it is not as significant as previously suggested due to the increase in radiative feedback experienced by galaxies. Much of the work in this chapter is to be submitted for publication in Sullivan, D. & Iliev, I. T (in prep.).

4.1 Introduction

Stars are considered the most likely sources of ionizing radiation during the EoR. Other possible sources, such as QSOs, although potentially very bright, are thought to be too rare to have a significant impact (e.g Madau et al., 1999; Barkana and Loeb, 2001; Faucher-

Giguère et al., 2008b,a; Willott et al., 2010; Kuhlen and Faucher-Giguère, 2012; Fontanot et al., 2012; Grissom et al., 2014; Haardt and Salvaterra, 2015). The ability for stars to drive cosmic reionization however depends on their ability to efficiently escape photons from their host galaxies. The escape fraction ($f_{\rm esc}$) of hydrogen ionizing photons from high-redshift galaxies is one of the most important parameters in any model of reionization, however is very poorly understood. In the concordance ΛCDM model, dwarf galaxies are thought to form early (e.g Somerville et al., 2003) and therefore dominate the hydrogen ionizing photon budget at the beginning of the EoR. In this scenario, semi-analytical models (e.g Wyithe and Cen, 2007; Shull et al., 2012; Robertson et al., 2013) require that $f_{\rm esc} > 10\%$ to reionize the Universe by $z \sim 6$ (as constrained via the spectra of high-redshift quasars, e.g Becker et al., 2001; Fan et al., 2001). This value is assuming that stars are not solely responsible for the reionization process, with QSOs also contributing.

Such a high $f_{\rm esc}$ is problematic in terms of observations, as with the exception of two objects (Leitet et al., 2011, 2013) no escaping ionizing radiation has been directly detected from the local Universe up to $z \sim 1$ (Leitherer et al., 1995; Deharveng et al., 2001). Indirect evidence exists for a few objects at $z \lesssim 1$ (e.g Heckman et al., 2011; Zastrow et al., 2011), although the lack of significant detections places upper limits of $f_{\rm esc} < 4\%$, however this could be due to detection bias (Bergvall et al., 2013). At $z \sim 3$, when the opacity of the IGM to ionizing photons is approximately less than unity, earlier reports of escaping Lyman continuum (LyC) photons from Lyman break galaxies (LBGs) and Lyα emitters (LAEs) have been marred by foregrounds (e.g Siana et al., 2015). Galaxy samples at this redshift find that $f_{\rm esc}$ is still well below the required values stated above, with 1σ upper limits of $f_{\rm esc} \sim 5\%$ (e.g Iwata et al., 2009; Boutsia et al., 2011). Attempts to constrain $f_{\rm esc}$ using numerical simulations have been met with difficulty, with predicted values ranging from $f_{\rm esc} < 10\%$ (Razoumov and Sommer-Larsen, 2006, 2007; Gnedin et al., 2008; Paardekooper et al., 2011; Kim et al., 2013; Kimm and Cen, 2014; Kimm et al., 2017; Trebitsch et al., 2017) and $f_{\rm esc} > 80\%$ (Wise and Cen, 2009; Razoumov and Sommer-Larsen, 2010). The high opacity to ionizing radiation during the EoR means we must rely on numerical models at $z > 3$, however clearly they predict vastly differing results. Another example of the discrepancy between these models lies with the dependence of $f_{\rm esc}$ on halo mass, with Gnedin et al. (2008) and Wise and Cen (2009) predicting that $f_{\rm esc}$ increases with halo mass, Kimm and Cen (2014); Wise et al. (2014); Paardekooper et al. (2015); Kimm et al. (2017) concluding the opposite, and Ma et al. (2015) finding no trend at all. While this is in part due to the variety of different codes and sub-grid prescriptions used in these works,

one of the difficulties in comparing individual models is that there is no single definition for $f_{\rm esc}$; such as the radius or the photon wavelength it is measured at.

The mechanisms which enable the escape of ionizing photons are complex, with Paardekooper et al. (2015) concluding that when $f_{\rm esc}$ is low, most photons are absorbed locally, within 100 pc of their source. For photons to escape, these absorbers must therefore be disrupted. Geen et al. (2015) showed that depending on the structure of molecular clouds and the strength of the source, the resulting H II regions may or may not be trapped entirely by the birth cloud. Therefore, it is unlikely that ionizing radiation alone can carve out channels in the gas through which to escape. Some studies have concluded that $f_{\rm esc}$ is regulated by SNe driven winds (e.g Kimm and Cen, 2014; Kimm et al., 2017; Trebitsch et al., 2017), allowing it to reach the required $f_{\rm esc} \sim 10\%$ to reionize the Universe. Interestingly, Ma et al. (2016) recently argued that binary stars could provide the photons required to complete reionization by this time with galaxies alone, using semi-numerical techniques. This would require $f_{\rm esc} \gtrsim 20\%$, a figure which models of reionization suggest is required to match the Thompson scattering optical depth to reionization (e.g Kuhlen and Faucher-Giguère, 2012; Finkelstein et al., 2012; Robertson et al., 2013, 2015). However, in their work radiative transfer is carried out in post-processing, therefore neglecting many of the negative self-feedback effects discussed in chapter 3. In this chapter, we probe whether feedback effects offset the predicted boost in $f_{\rm esc}$ when including a stellar evolution model which includes a physically and observationally motivated treatment of binary stars. We carry out two fully-coupled radiation hydrodynamical simulations of the EoR to determine how binaries impact $f_{\rm esc}$, using the single and binary evolutionary models of the Binary Populations and Stellar Populations Synthesis model (BPASS; Eldridge et al., 2008).

This Chapter is arranged as follows: in section 4.2 we present our choice of simulation code, basic simulation results, and methodology for measuring $f_{\rm esc}$ self-consistently from our simulations. In section 4.4 we present our predictions for the instantaneous and time averaged $f_{\rm esc}$, as a function of mass for our binary and single stellar evolutionary models. Finally, in sections 4.5 and 4.6 we discuss our findings in the context of the literature and provide concluding remarks.

4.2 Method

4.2.1 Radiation Hydrodynamics Code

To probe the impact of stellar population uncertainty on the escape fraction of ionizing photons in cosmological simulations, we use the Eulerian adaptive mesh refinement (AMR) code RAMSES (Teyssier, 2002, ver. 3). The code is based on the fully threaded oct-tree structure of (Khokhlov, 1998), with grid refinement performed on a cell-by-cell basis, following user desired criteria. A second-order Godunov scheme, proven to capture shocks and follow the thermal history of the fluid accurately, is used to solve the equations of hydrodynamics in their conservative form. These states are reconstructed on cell interfaces using a MinMod method and advanced using the Harten-Lax-van-Leer contact wave Riemann solver (HLLC; Toro et al. 1994). Dark matter and stars are included in the form of a collisionless system of N-body particles, described by the Vlasov-Poisson equations. Gas is allowed to cool through atomic excitation/de-excitation down to 10^4 K and further to ~ 100 K via metal fine-line transitions.

We use the multi-group RT module RAMSES-RT, developed by Rosdahl et al. (2013), to compute the photoionization of three species: HI, HeI, and HeII. We adopt the *on-the-spot* approximation, where UV photons emitted from recombinations are assumed to ionize an atom in the same grid cell, hence case-B recombination rates are used when computing the gas cooling rate (Hui and Gnedin, 1997). The star-formation and SNe feedback models and the values of code specific parameters are calibrated and set following the methods set out in section 2.7.

4.2.2 Binary Populations and Stellar Population Synthesis

As introduced in section 1.6.3, stellar population synthesis (SPS) models are tools for interpreting the integrated light observed from galaxies, based on a theoretical understanding of the properties of individual stellar populations. One key detail in the majority of SPS models, is their inability to accurately track the evolution of interacting binaries. The Binary Population and Spectral Synthesis code, BPASS[1] (Eldridge et al., 2008), aims to address these uncertainties by using a large grid of 250,000 stellar models to accurately follow the evolution of interacting binaries. They find that including phenomena such as mass loss and transfer in binary systems improves agreement with observations of local stellar populations, as it increases the probability that stars lose their hydrogen

[1] http://bpass.auckland.ac.nz

envelope thus leading to fewer red supergiants, more Wolf–Rayet stars, and more Type Ib/c SNe than single-star models. Furthermore, binary stars leads to harder UV spectrum (increased luminosity at low-wavelengths), which is needed to explain nebula emission line properties observed in young metal-poor galaxies at $z \sim 2-3$, and local galaxies (Steidel et al., 2014). The hardening of UV emission is most prominent for old stellar populations, as mass transfer leads to an increase in abundance of hot, massive stars, long after the initial burst of star formation. Additional to the increase in abundance of old massive stars, efficient mass transfer can also spin up secondary stars to high rotation rates, leading to strong mixing in metal poor main-sequence stars which do not generate strong stellar winds (Cantiello et al., 2007; Eldridge and Stanway, 2012). Material produced in the stellar core is then immediately redistributed in the envelope, leading to hotter (bluer) stars on average in old stellar populations.

The above physical phenomena leads to a significant enhancement in UV flux for the wavelength range spanning hydrogen and helium ionization, as shown in figure 4.1 for the binary (solid) and single (dot-dashed) BPASS models. Here we show the SED for stellar populations at solar metallicity and ages 1 Myr (red), 10 Myr (blue), and 19 Myr (orange). As shown in Ma et al. (2016), after $t \sim 3$ Myr roughly 60 % of ionizing photons have been emitted in the binary model, compared to only approximately 40 % in the single population model at low metallicity ($Z = 0.05\ Z_{\odot}$; where $Z_{\odot} = 0.02$). The discrepancy between the two models is lessened towards higher metallicity, however a small boost in ionizing photon production is still seen at $Z_{\odot}$.

4.3 Cosmological Simulations

We assume a geometrically flat ΛCDM Universe consistent with a WMAP-7 cosmology (Komatsu et al., 2011) with cosmological parameters: $\Omega_{\rm M} = 0.272$, $\Omega_{\Lambda} = \Lambda_0/(3H_0^2) = 0.728$, $\Omega_{\rm b} = 0.045$, $h \equiv H_0/(100\ {\rm km\ s^{-1}\ Mpc^{-1}}) = 0.702$, and $\sigma_8 = 0.82$, where the symbols take their usual meaning. Initial conditions are generated using the MUSIC code (Hahn and Abel, 2013) at redshift $z = 100$, in a comoving volume of ~ 5.7 Mpc on-a-side. Variations in these cosmological parameters are not expected to largely influence our results.

The coarse grid level is set to $\ell_{\rm min} = 8$ in the AMR hierarchy for all simulations, resulting in $(2^{\ell_{\rm min}})^3 = (256)^3$ DM particles and coarse grid cells, such that our DM particle mass $M_{\rm DM} \sim 3.4 \times 10^5\ {\rm M}_{\odot}$. We set the maximum level of refinement $\ell_{\rm max} = 18$, such that our effective spatial resolution is $\Delta x_{\rm min} = 21.72$ pc comoving and our stellar particle mass

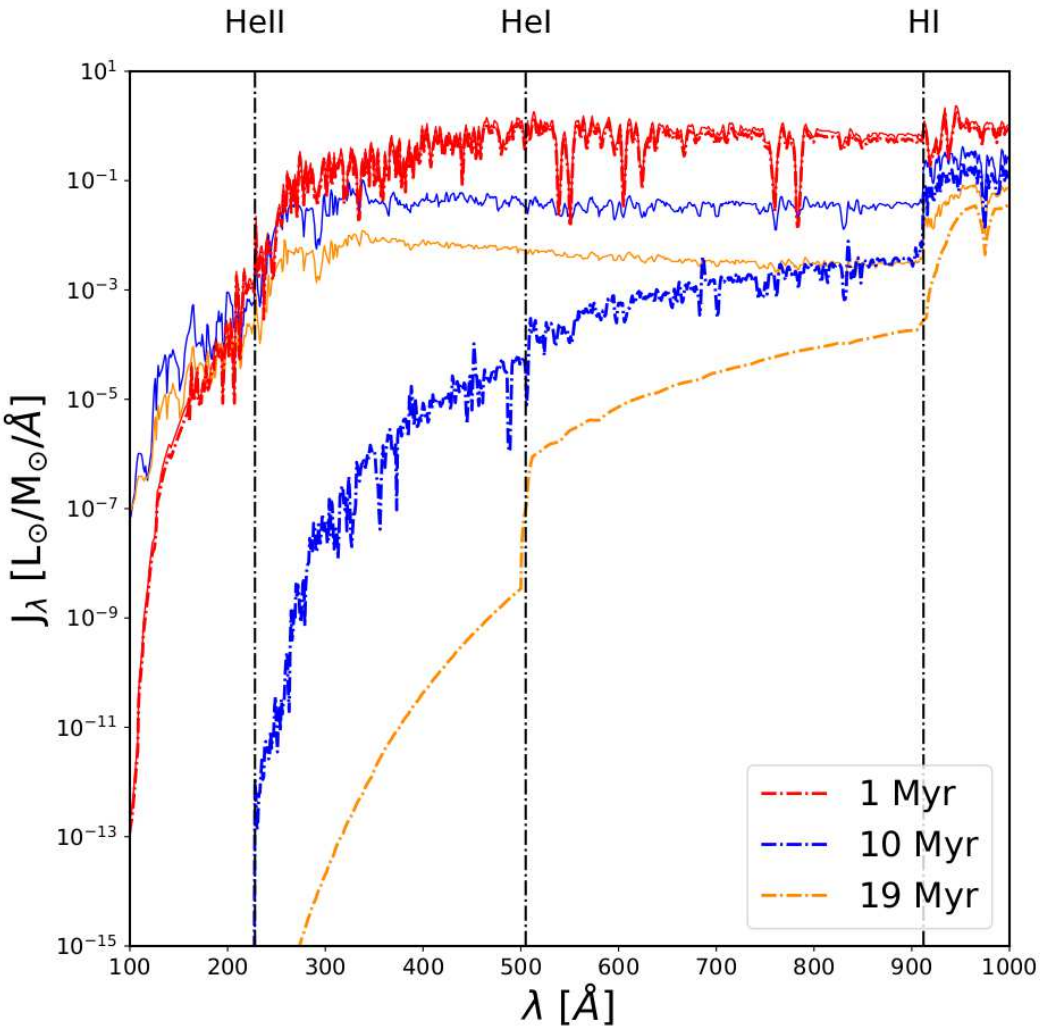

Figure 4.1: Spectral Energy Distributions (SEDs) for the BPASS binary and single stellar population models (solid and dot-dashed, respectively), at ages 1 Myr, 10 Myr, and 19 Myr and solar metallicity ($Z = Z_\odot$). The vertical lines show the ionization wavelengths for HI, HeI and HeII. For young stellar populations (red), there is very little difference between the two models. However, old binary populations have a much harder UV spectrum (larger flux at low wavelength), due to mass transfer and other evolutionary effects discussed in the text.

is $m_{*,\rm min} \sim 2 \times 10^3 M_\odot$. We adopt uni-grid initial conditions for the gas and dark matter (i.e, no zoom regions).

We calibrate our simulations and adopt star formation and SNe feedback parameters based on the methodology outlined in sections 2.7.1 and 2.7.2, with:

$$\begin{aligned} n_* &\geq 8 \text{ cm}^{-3} \\ T_0 &= 540 \text{ K}, \\ g_* &= 2, \end{aligned} \tag{4.1}$$

where g_* is the polytropic index. This ensures the Jeans length of star-forming gas is always resolved by at least six cells in our simulations. Due to computational constraints, we adopt a reduced speed of light, $f_{\rm c} = 1/100$, which relaxes the Courant-condition for the radiation fluid, reducing the number of radiation sub-cycles necessary to reach thermal convergence (Gnedin and Abel, 2001).

Dark matter (sub)-haloes are catalogued using the ROCKSTAR phase-space halo finder (Behroozi et al., 2013b), while merger trees are constructed using the CONSISTENT-TREES algorithm (Behroozi et al., 2013c). Particles are first divided into 3D Friends-of-Friends (FOF) groups, which are analysed in 6D phase-space to give robust, grid- and shape-independent haloes. Haloes that are not gravitationally bound are removed and properties are computed using the virial overdensity definition of Bryan and Norman (1998).

Photoionization of both hydrogen and helium is followed via the coupling of UV photons with the gas, we use the above binary and single stellar population models of Eldridge et al. (2008) assuming a Kroupa IMF (Kroupa, 2002), from which we tabulate SEDs for a variety of ages and metallicities as shown in figure 4.1. Furthermore, we include the homogeneous redshift evolving UV background of Faucher-Giguère et al. (2010) to account for UV emission from non-stellar sources, such as quasars, as shown in the right-hand panel of figure 2.3. In figure 4.2, we show the reionization histories (neutral fraction, volume and mass weighted), photoionization rate (volume and mass weighted), and the integrated Thompson scattering optical depth of both our binary and single stellar population models as a function of redshift. While both models undergo reionization early, with respect to constraints on the neutral fraction, we note that both models are consistent with the Planck Collaboration et al. (2016b) constraints on the optical depth, while only the binary model shows minor tension with the more recent Planck Collaboration et al. (2016a) predictions. In section 3.3 we discussed in depth the variation in reionization histories exhibited in small-volumes, such as ours. As we are not concerned

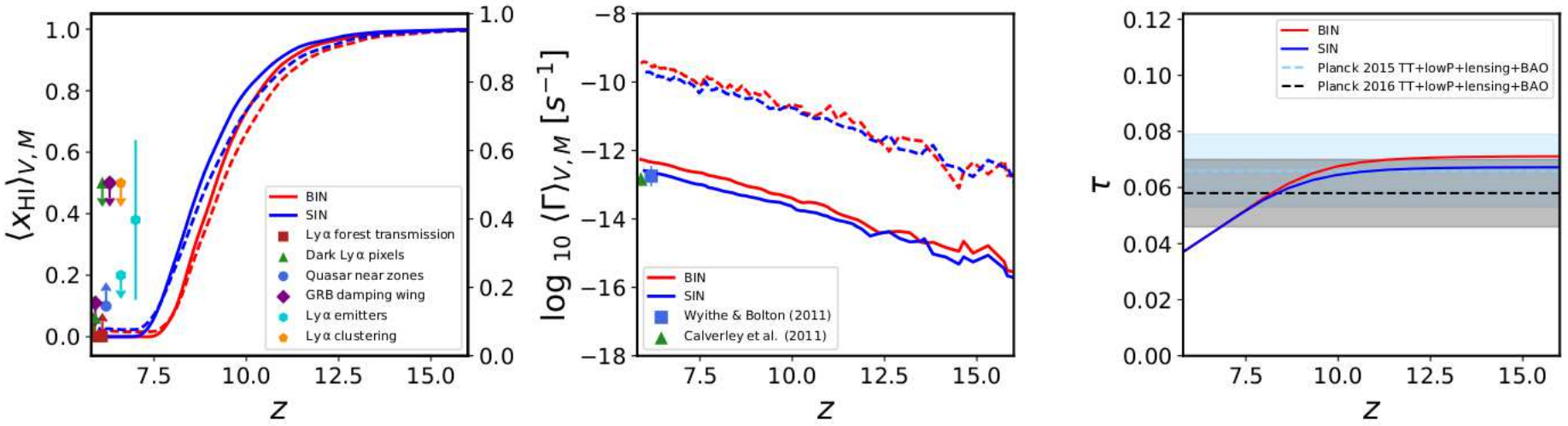

Figure 4.2: *Top panel:* the volume/mass-weighted (solid/dahsed respectively) mean neutral fraction of hydrogen for the binary (BIN) and single (SIN) stellar population models, compared to observational constraints from Ly α forest transmission (squares, red; Fan et al., 2006), dark Ly α forest pixels (triangles, green; McGreer et al., 2011, 2015), quasar near zones (circles, blue; Schroeder et al., 2013), GRB damping wind absorption (diamonds, voilet; McQuinn et al., 2008; Chornock et al., 2013), and decline in Ly α emitters (hexagons, cyan; Ota et al., 2008; Ouchi et al., 2010) following from Robertson et al. (2015). *Middle panel:* The mean volume/mass-weighted photoionization rate, compared to the volume-weighted observations of Calverley et al. (2011) and Wyithe and Bolton (2011) as the (green) triangle and (blue) square, respectively. *Bottom panel:* The integrated electron-scattering optical depth compared to the PlanckTT+lowP+lensing+BAO 2016 (thin black horizontal line) and 2015 results (thin lightblue line), with 1σ error intervals shown by the shaded regions (Planck Collaboration et al., 2016a,b).

with the global properties of reionization in this work, it should not have a significant impact on our results.

Our simulations, like those presented in chapter 3, exhibit an inside-out reionization where the mass-weighted neutral fraction is consistently below the volume-weighted quantity during the EoR, which is characteristic of an inside-out progression. The initial evolution of both the volume and mass weighted neutral fractions is steeper when binaries are included, as they emit a larger fraction of their cumulative photons produced in the first 3 Myr of their lifetime than single stellar populations. As a result, the initial Strömgren radius, r_S, is larger as:[2]

$$r_S \approx \left(\frac{3}{4\pi} \frac{\dot{N}_{\rm ion}}{n_{\rm H}^2 \alpha_{\rm B}(T)} \right), \tag{4.2}$$

where $\dot{N}_{\rm ion}$ is the photon production rate (photons per second), $n_{\rm H}$ is the number density of hydrogen, and $\alpha_{\rm B}$ is the case-B recombination rate (Hui and Gnedin, 1997). The initial growth of H II regions is therefore faster when binary populations are acounted for, which combined with the inclusion of the UV background model of Faucher-Giguère et al. (2010) (to account for non-stellar sources) leads to an earlier reionization in these simulations when compared with those in the previous chapter. Furthermore, the harder UV spectrum of old binary populations could lead to an increased heating rate, especially in dense cores which are otherwise difficult for ionizing radiation to penetrate.

4.3.1 Estimating the Escape Fraction

To compute the ionizing escape fraction on a halo-by-halo basis, we follow the methodology of Kimm and Cen (2014) by using the HEALPIX algorithm (Górski et al., 2005) to calculate the net ionizing flux along rays cast radially from the centre out to the virial radius, $R_{\rm vir}$. This leads to the assumption that each halo be a point source, such that photons are assumed to arrive at the virial radius at time $(t - R_{\rm vir})/c'$ after being emitted from the halo centre, where c' is the reduced speed of light adopted. The instantaneous escape fraction is then given by:

$$f_{\rm esc}(t) = \frac{\int d\boldsymbol{\Omega}\ \vec{F}_{\rm ion}(t) \cdot \hat{r} \Theta(\vec{F}_{\rm ion} \cdot \hat{r})}{\int dm_* \dot{N}_{\rm ion}(t - R_{\rm vir}/c)}, \tag{4.3}$$

[2]Strictly this Strömgren radius is only valid for a homogeneous medium with a static source, and assumes no back reaction of the radiation to the gas. This value therefore should be interpreted as an approximation only.

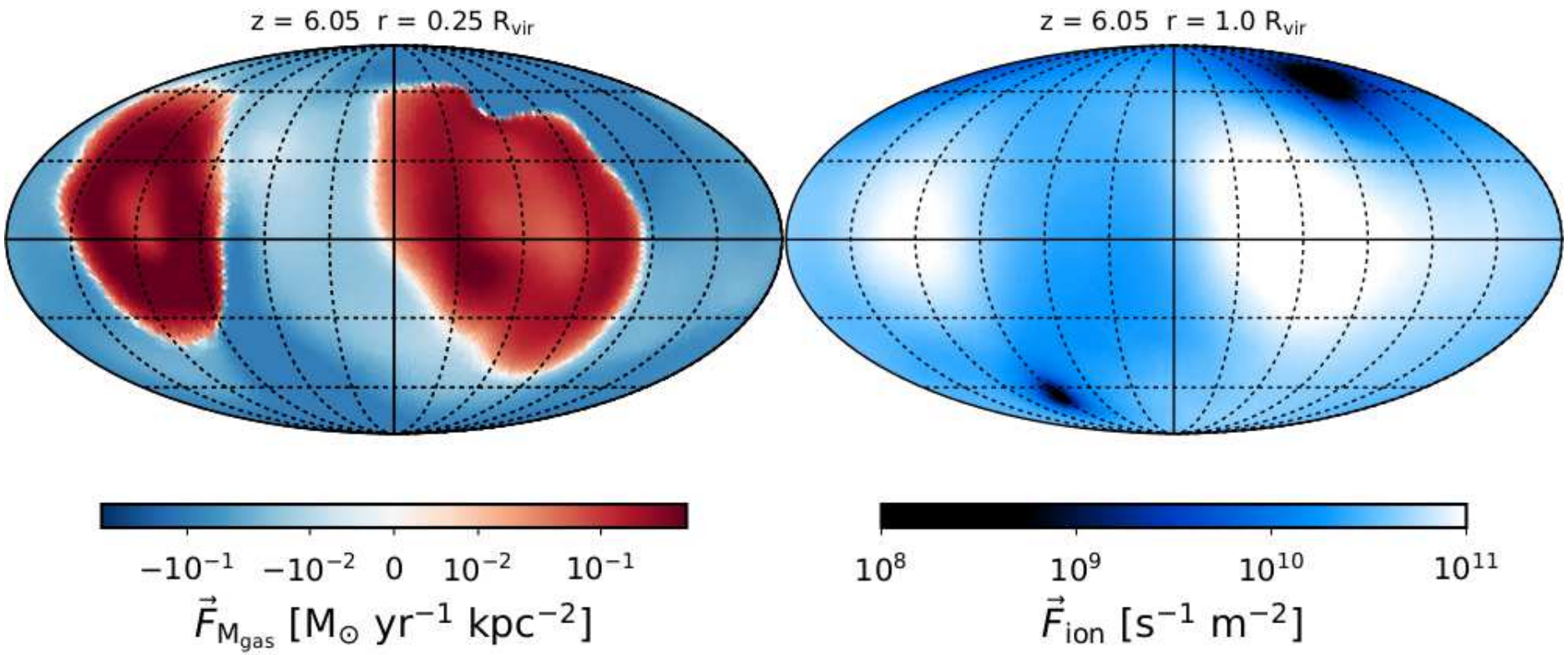

Figure 4.3: Mollweide view of the flux of mass and ionizing photons (left and right, respectively) on the sky for a halo chosen at random, where the radius and redshift at which the fields are projected on the sky is indicated above the panels in units of the halo virial radius. There is a clear correlation on the sky between strong outflow or inflow, with regions of high or low ionizing flux, respectively, suggesting that SNe driven outflows could be responsible for regulating $f_{\rm esc}$.

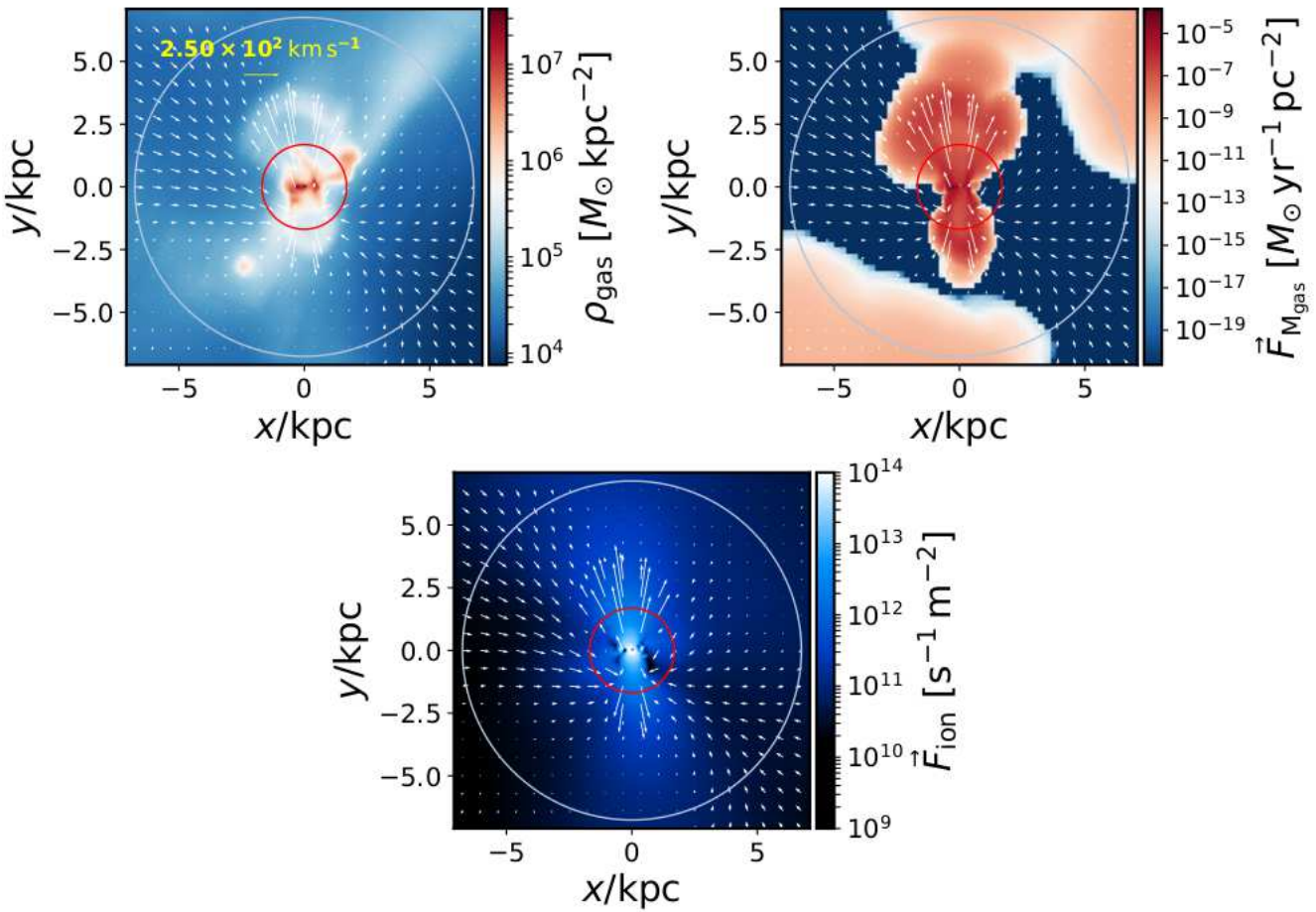

Figure 4.4: *Top left:* projection of gas density, rotated side on to the stellar disc of the central galaxy, for the same halo used to produce figure 4.3. The lightblue and red annotated circles show the virial radius and 1/4 of its value, the points at which the ionizing and mass fluxes are projected onto the sky, respectively. The projection shows a strong bi-polar blast wave emanating from the disc, sweeping up material in its path and carving a hollow channel through the halo. *Top right:* Slice through the centre of the halo, highlighting the bi-polar outflow seen in the top left panel (red regions). *Bottom panel:* Slice through the ionizing photon flux, through the centre of the halo. Photons primarily escape along the outflow trajectory, through underdense channels. The quiver arrows annotate a slice through the gas velocity field and is the same in all panels.

where $d\mathbf{\Omega}$ is the solid angle, $\vec{F}_{\rm ion}$ is the ionizing photon flux, $\hat{r}$ is the radial unit vector, $\dot{N}_{\rm ion}(t)$ is the ionizing photon production rate per unit solar mass of a simple stellar population at time t, and Θ is the Heaviside step function:

$$\Theta(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2} & x = 0 \\ 1 & x > 0 \end{cases}$$

which ensures we only consider photons travelling radially outwards. While this method may not be the most robust, Trebitsch et al. (2017) recently compared a similarly motivated method with a more robust calculation, whereby they compute the escape fraction through ray-tracing for each star particle. By doing so, they do away with the assumption that haloes are point sources and accurately account for the light travel time across the halo. Furthermore, the M1 moment method is known to be too diffusive (Iliev et al., 2006a, 2009; Rosdahl et al., 2013), which could lead to leakage around absorbers. However, Trebitsch et al. (2017) conclude that both methods show remarkably good agreement with each other, largely due to the fact that stars are highly concentrated around the halo centre which alleviates both of the issues raised above. One issue which could severely impact the escape fraction, which we have not considered here, is the dust content of galaxies. As we are concerned with high-redshifts only, where the gas has not been significantly enriched with metals and thus expected to have low dust growth (Lisenfeld and Ferrara, 1998; Engelbracht et al., 2008; Galametz et al., 2011; Fisher et al., 2014), we conclude that dust is unlikely to have a significant impact on the escape fraction in agreement with Kimm and Cen (2014).

As previously discussed, SNe driven winds are thought to be the primary mechanism for regulating $f_{\rm esc}$. Figure 4.3 shows the flux of (gas) mass and ionizing photons (left and right, respectively) on the sky for a halo chosen at random, where the flux of mass is computed as:

$$\vec{F}_{\rm M_{gas}} = \int d\mathbf{\Omega}\ \rho\vec{u}\cdot\hat{r} = \vec{F}_{+} + \vec{F}_{-}, \tag{4.4}$$

where ρ and $\vec{u}$ denote the gas density and velocity respectively, and we have split the total flux into positive and negative components which correspond to outflow and inflow.

The correlation of peaks on the sky suggests that the escape of photons here is highly regulated by strong SNe driven outflows, which carve out cavities in the halo through which photons can escape, as discussed above. Furthermore, we find minima in the photon flux

on the sky where there are net inflows, as one might expect, although the photon flux is more extended on the sky. The reason for this is twofold: first, the photon flux is measured at the virial radius while the outflow rate is measured at 1/4 of its value, therefore opening up a larger angle on the sky, and secondly because the M1 moment method is known to be quite diffusive and struggles to properly capture shadows behind absorbers (Iliev et al., 2006a, 2009; Rosdahl et al., 2013). We choose to measure the outflow rate at 1/4 of the virial radius rather than the full radius, as past this point the halo is almost entirely made up of diffuse gas supported by a hot atmosphere. Photons which are able to reach this part of the halo almost certainly continue to escape through the virial sphere, as the vast majority of absorbers exist locally to the ionizing source as discussed above. Therefore, if strong outflows are generated near the central galaxy, we can expect a subsequent peak in the flux of photons across the virial radius (albeit with some delay due to the finite light travel time).

To demonstrate the accuracy of our method, we show a projection of the gas density (top left), and slices of the outflow rate (top right) and photon flux (bottom centre) through the centre of the halo chosen for the above mollweide views, in figure 4.4. In each panel, we annotate the gas velocity field to show the direction of flow, where the scale-bar in the top left panel applies throughout. The lightblue and red annotated circles denote the halo virial radius and 1/4 of its value, respectively, which are the respective radii where we compute the photon and mass flux on the sky. All panels are orientated such that the stellar disc of the galaxy is side-on. In the top left panel, a bipolar SNe driven blast wave can be clearly seen above and below the disc, where the shock front has swept up material leaving an underdense channel. This is highlighted in the right hand panel, where the velocity vectors clearly show the trajectory of the material. Finally, in the bottom panel, we find that the corresponding low density channel allows photons to flood out of the galaxy, ultimately escaping the virial sphere.

4.4 Results

4.4.1 Instantaneous and Time Integrated Escape Fraction

As pointed out by Kimm and Cen (2014), cosmological hydrodynamical simulations with SNe feedback often suffer from the artificial over-cooling problem when forming disc galaxies (e.g Creasey et al., 2011; Hummels and Bryan, 2012). The net effect is a weakening of feedback, as most of the energy released in the SNe event is injected as thermal energy,

which the gas can radiate away efficiently before it is converted into momentum. This is particularly problematic when the ISM is inadequately resolved, severely inhibiting the ability to efficiently escape photons from the central disc. Kimm and Cen (2014) overcame this challenge by implementing a mechanical SNe feedback scheme, which accurately approximates the Sedov blast wave solution from the free expansion to snowplow stages. This scheme recovers the analytical solution of Thornton et al. (1998) at low resolutions where traditional thermal dump models fail (see figure 16 of Kimm and Cen, 2014).

At the time of writing, this mechanical feedback model is not publicly available, therefore we adopt the kinetic feedback model of Dubois and Teyssier (2008) as described in section 2.7.3. While we note that this scheme is not immune to the over-cooling problem, it better captures the Sedov blast wave solution over traditional thermal dump schemes, as energy is injected into the grid as kinetic energy. In figure 4.5, we present the instantaneous escape fraction (top row) and number of ionizing photons escaping the virial sphere (bottom row) for the binary and single stellar population models at refshifts $z = 6$ (left column), $z = 8$ (middle column), and $z = 9$ (right column) as a function of halo virial mass. The points in both rows show the scatter of the data, which the error bars show the mean and standard deviation in each mass bin. A small percentage of haloes, $\sim 10\%$, at each redshift have $f_{\rm esc} > 1$. Upon individual inspection, this is typically due to a nearby massive neighbour, from which ionizing photons stream through the halo under consideration. This boosts our estimated value, and when this is the case, we follow Kimm and Cen (2014) in randomly reassigning $f_{\rm esc}$ between $0.9 - 1.0$.

We find that $f_{\rm esc}$ decreases with increasing halo mass, in good agreement with Kimm and Cen (2014); Wise et al. (2014); Paardekooper et al. (2015), however shows very little redshift evolution. At $z = 6$ and $z = 8$, $f_{\rm esc}$ shows minor enhancement in the binary population model at high masses, however, is consistent with the single population model within 1σ. A similar trend is observed for the number of ionizing photons which escape the virial sphere, however this is proportional to halo mass (as one might expect). Again, the mean is slightly raised for the binary model, however, not significantly. The bottom row of panels show good agreement with the predictions of Kimm and Cen (2014) at all masses, which is not of great surprise as we largely follow their methodology.

The instantaneous escape fraction however is a poorly defined and unstable quantity. For instance, short lived star bursts, which do not represent the overall star formation history of the halo, can strongly influence the estimated value of $f_{\rm esc}$. Therefore, we

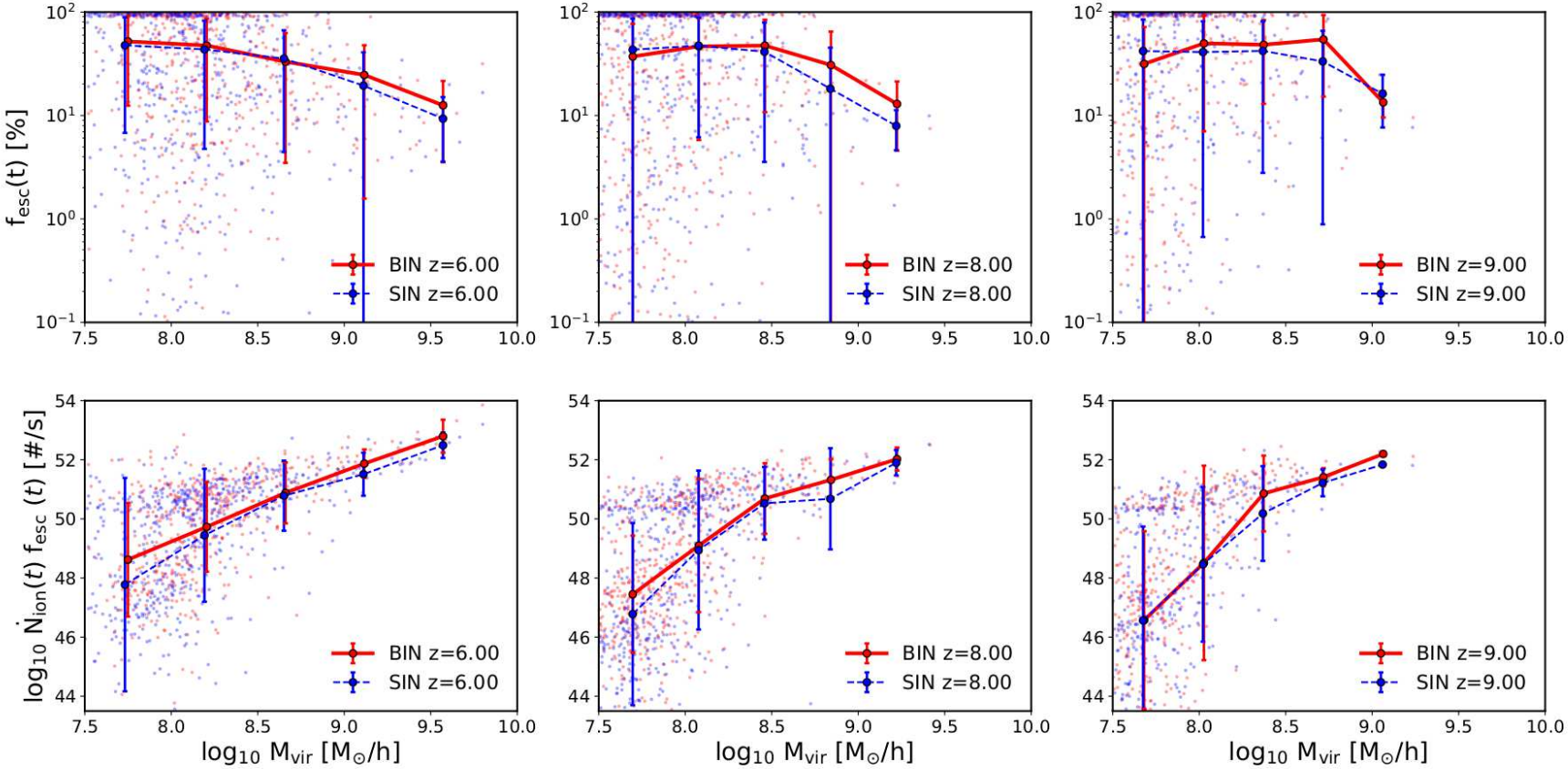

Figure 4.5: *Top row, left to right:* Instantaneous escape fractions of halos as a function of their virial mass at redshifts $z = 6$, $z = 8$, and $z = 9$, where the points show the scatter in the data and the error bars show the mean and standard deviation in each mass bin for binary (red) and single (blue) stellar populations. The escape fraction evolves strongly with halo mass, however exhibits very little evolution with redshift. *Bottom panel, left to right:* Instantaneous number of ionizing photons escaping the virial sphere as a function of halo mass, and for the same redshifts shown in the upper panel.

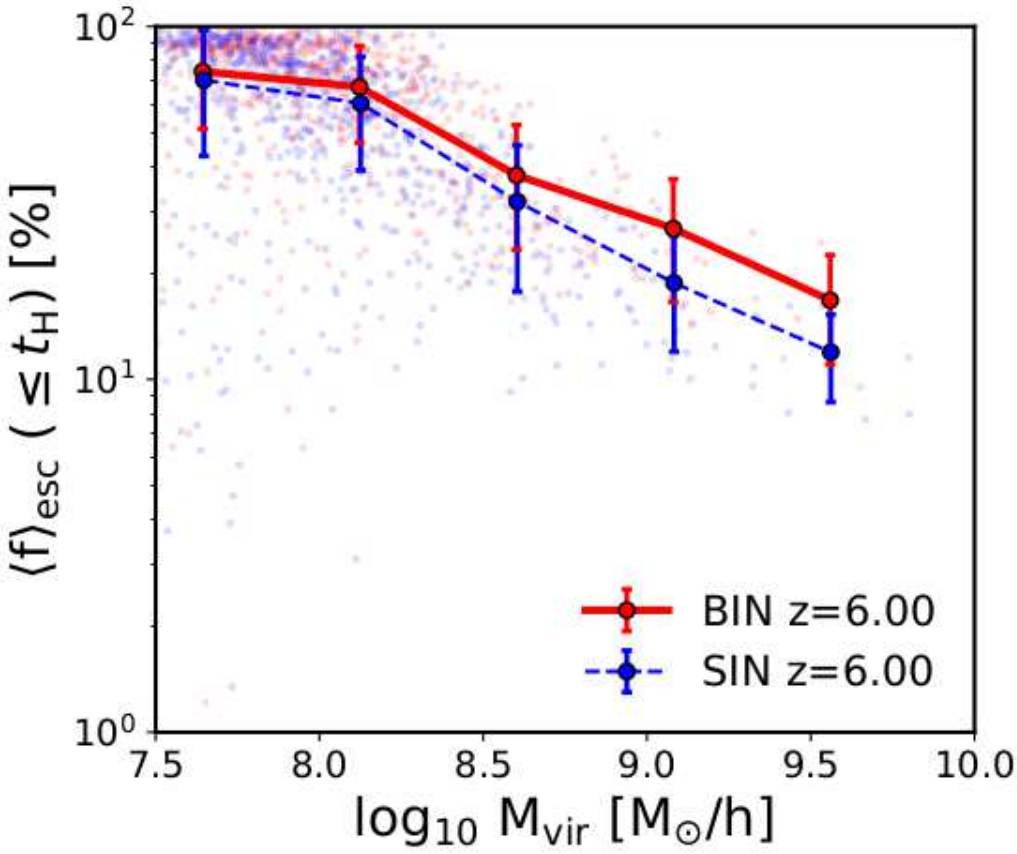

Figure 4.6: Photon-production rate weighted time integral of the escape fraction of ionizing photons, as a function of the final halo mass (measured at time $t = 0$). This quantity is more robust than the instantaneous value, as it factors in the full star formation history of the halo (accounting for short lived star bursts). While the binary model (red) on average results in larger escape fractions, we note that it is not significantly enhanced over the single stellar population model (blue).

compute the photon-production rate time-averaged escape fraction as:

$$\langle f_{\rm esc}\rangle(\leq t_{\rm H}) = \frac{\int_0^{t_{\rm H}} \dot{N}_{\rm ion}(t) f_{\rm esc}(t) dt}{\int_0^{t_{\rm H}} \dot{N}_{\rm ion}(t) dt}, \tag{4.5}$$

where $t_{\rm H}$ denotes the age of the halo over which the integration is performed. To do this, we descend the halo merger tree, produced using the CONSISTENT TREES algorithm (Behroozi et al., 2013c), following the most massive progenitor line. The time-integrated escape fraction is shown in figure 4.6, as a function of the final halo virial mass. Again, we find that the escape fraction both evolves strongly with mass and is enhanced due to binary evolutionary effects, however no haloes exhibit $\langle f_{\rm esc}\rangle(\leq t_{\rm H}) > 1$. Linear extrapolation of both the binary and single stellar fits reveal escape fractions of 1.55% and 1.20%, respectively, for a Milky Way sized halo with virial mass $M_{\rm vir} \sim 1 \times 10^{12}\ M_{\odot}$. This is consistent with observations of the local Universe (e.g Leitherer et al., 1995; Deharveng et al., 2001; Heckman et al., 2011; Zastrow et al., 2011).

4.4.2 Jeans Filtering And Stellar Abundance

The hardening of the UV spectrum of SSPs due to binary populations (see figure 4.1) suggests that there may be additional heating during reionization, as these photons could penetrate deeper into cold neutral gas. In section 1.3.3, we derived the cosmological analogue to the Jeans mass scale, known as the *filtering mass*. This is the mass scale over which baryonic perturbations are smoothed with respect to dark matter, and accounts for the entire evolution of the Jeans mass over cosmic time. It is therefore sensitive to the full thermal history of the gas, as the Jeans scale depends on the adiabatic sound speed via equations 1.17 and 1.19. Any net increase in the gas temperature over the single stellar population model could potentially lead to a rise in the filtering mass, thus suppressing the growth of structure by raising the minimum mass which is unstable to collapse. While this is assuming that a linear analysis holds, which in truth is not correct for reasons discussed in sections 1.3.5 and further demonstrated by our measurements of the characteristic mass scale in chapter 3, it would in principle indicate whether simple stellar evolution uncertainties could suppress the growth of structure. We compute the filtering and Jeans mass, assuming matter domination using equation 1.20 and evaluated at the mean density of the Universe, as shown in figure 4.7 where the solid lines show the filtering mass, M_F, and the dashed lines the Jeans mass, $M_{\rm J}$.

We show the complete evolution of both the Jeans mass and filtering mass to demonstrate their likeness to the $v_{bc} = 0$ case from figure 1.4, but zoom in the range $6 \leq z \leq 10$ to show the differences between both models. Increased heating due to the hardening of

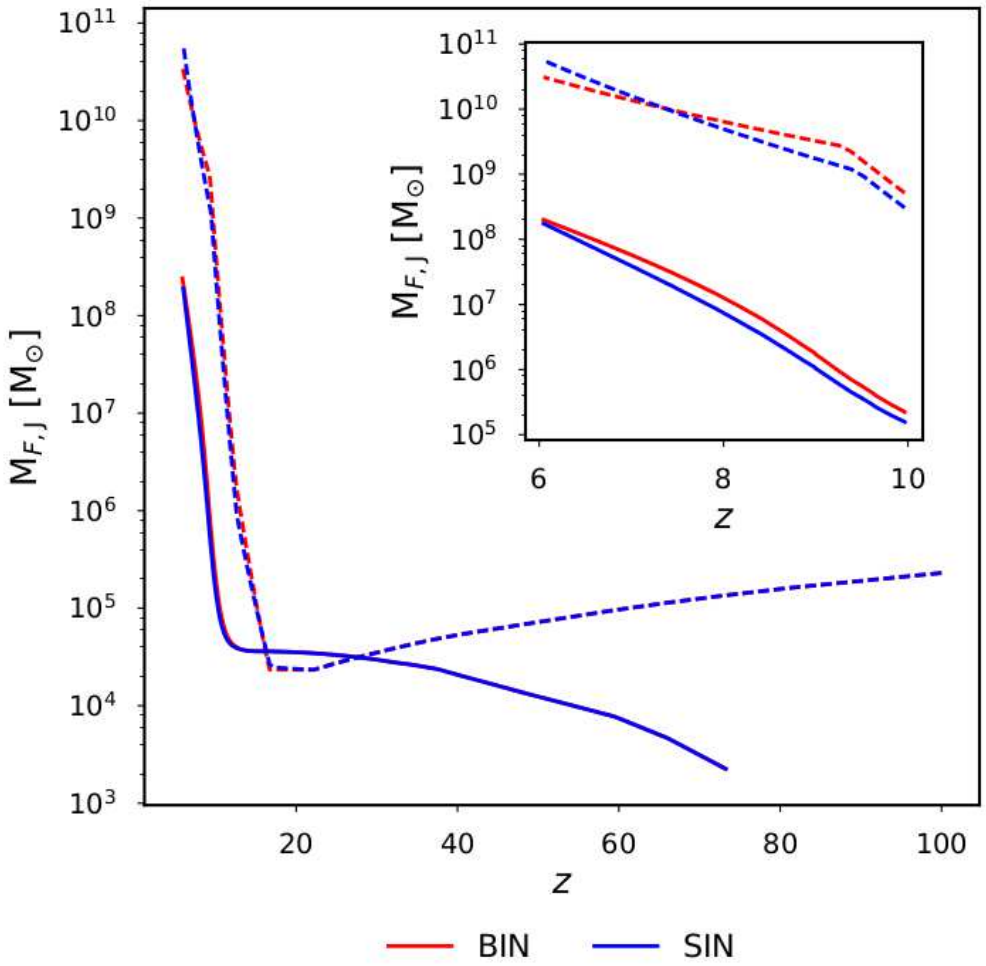

Figure 4.7: Evolution of the filtering mass, M_F (solid lines), and Jeans mass, M_J (dashed lines) evaluated at the mean density of the Universe, over the full redshift range of both our binary and single stellar population simulations (red and blue, respectively). The hardening of the UV spectrum due to binary population effects leads to additional heating, which in turn raises the adiabatic sound speed, c_s, and therefore both the Jeans and filtering masses. The filtering mass remains enhanced throughout reionization as it depends on the full thermal history of the gas.

the UV spectrum raises the adiabatic sound speed of the IGM, which in turn raises the Jeans mass and thus the filtering mass. While marginal, the effect is sustained throughout reionization despite the Jeans mass decreasing with respect to the single stellar population model towards $z \sim 6$, leading to a small (but significant) modulation of the stellar abundance of haloes, as shown in figure 4.8. Although both models are consistent within the 1σ error bars, any reduction in the mean stellar abundance could offset any potential boost in $f_{\rm esc}$ and subsequently the number of ionizing photons, which escape to the IGM as predicted by Ma et al. (2016).

The reduction in mean stellar abundance of low mass haloes, albeit minor, also impacts the faint end of the UV luminosity function, as shown in figure 4.9. While once again the discrepancy is small, the net result is that the single stellar population catches up with the binary case towards the end of reionization, meaning that any additional feedback in either model becomes self-regulating over time and thus less important with time. We compare our results with the observational constraints from Bouwens et al. (2015) and their resulting Schecter function fit (points and dotted lines, respectively), at redshifts $z = 6$, 7, and 8. Due to the limited volume of our simulations, our results match the observations only for the faintest data point at $z = 6$, and depart from the Schecter function fits at the faint end. While poorly constrained, there is limited evidence for the flattening out of the UV luminosity function towards the faint end (e.g Parsa et al., 2016).

4.4.3 Feedback-regulated Escape of Ionizing Photons

In figure 4.10, we show the time evolution of the star formation rate (blue shaded, left hand axis), outflow rate (orange dot-dashed, left hand axis), inflow rate (green dotted, left hand axis), instantaneous escape fraction, $f_{\rm esc}(t)$, (red line, right hand axis), and finally the photon-production weight time averaged escape fraction, $\langle f_{\rm esc} \rangle (\leq t_{\rm H})$ (purple dashed line, right hand axis), for two haloes with masses $M_{\rm vir} \simeq 6.29 \times 10^9\ M_\odot\ h^{-1}$ and $M_{\rm vir} \simeq 2.11 \times 10^9\ M_\odot\ h^{-1}$, respectively, from our binary population simulation. The bottom x-axis displays the time in Myr, while the upper x-axis shows the cosmological redshift. Both the outflow rate and $f_{\rm esc}(t)$ peak following an episode of star formation, due to the resulting peak in SNe activity which also acts to regulate the star formation rate. Although $f_{\rm esc}(t)$ fluctuates significantly in both haloes, $\langle f_{\rm esc} \rangle (\leq t_{\rm H})$ remains roughly constant in time.

We find that the escape fraction and star formation rate are anti-correlated, suggesting that photons are absorbed locally within the stellar birth cloud until disrupted by SNe

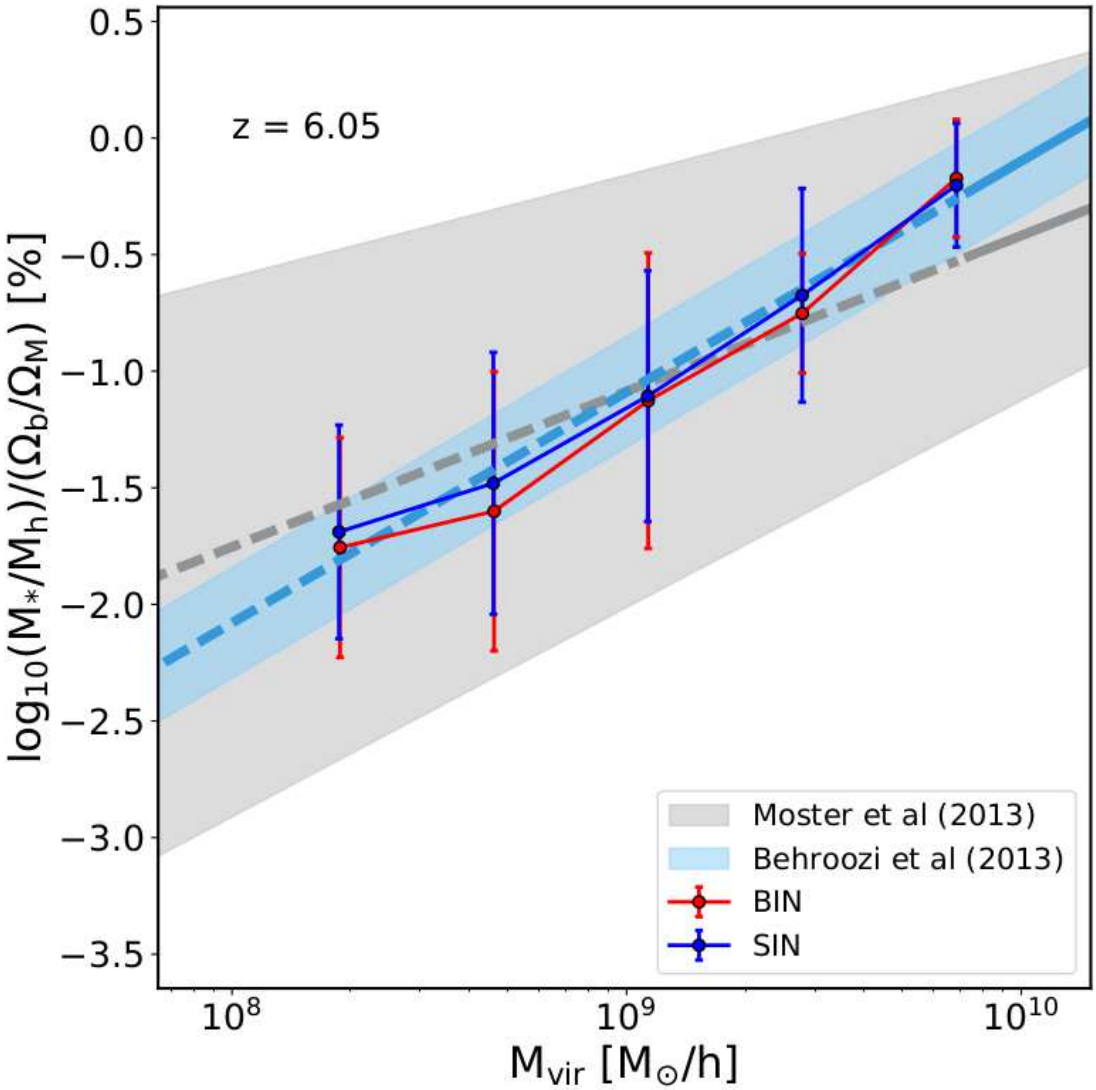

Figure 4.8: Stellar fraction of haloes (in units of the cosmic mean baryon fraction) for our binary and single stellar population models, compared with the models of Moster et al. (2013) and Behroozi et al. (2013a). The marginal, but significant, reduction in the mean stellar abundance of our binary model can be attributed to the raised filtering mass as a result of additional UV heating.

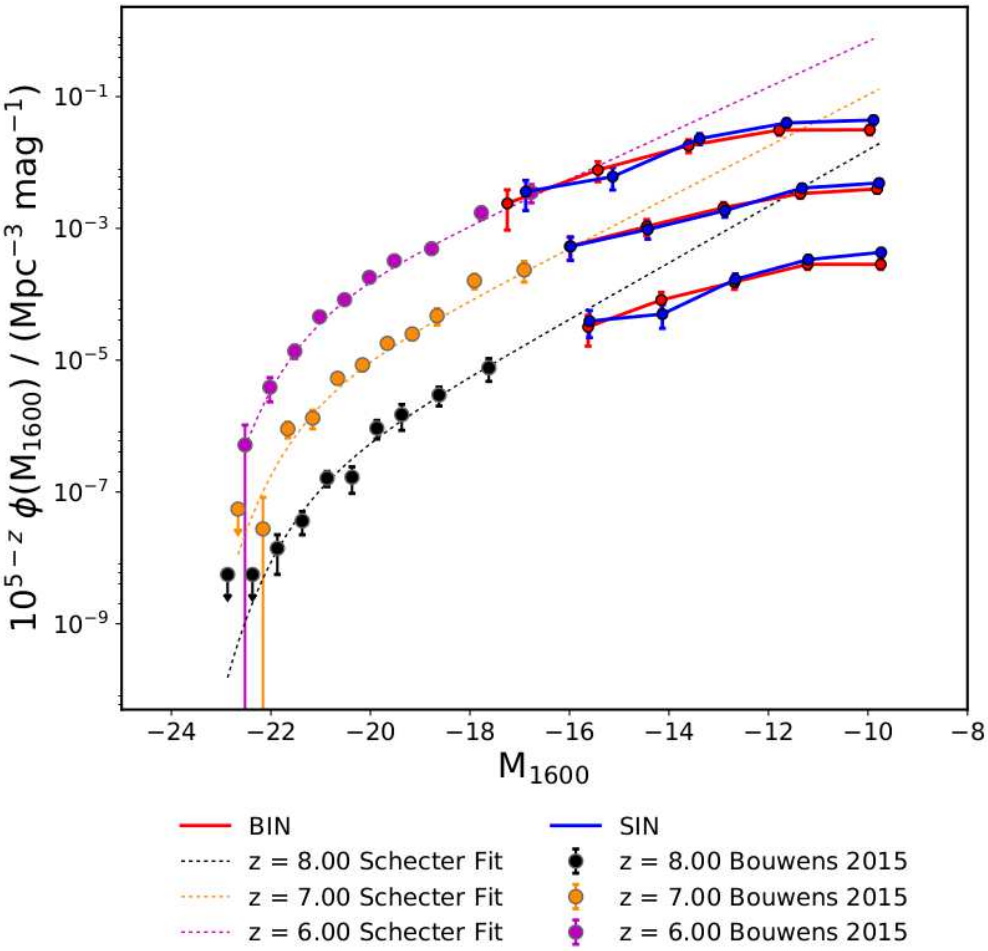

Figure 4.9: UV luminosity functions at $z = 6$, 7, and 8 for our simulated haloes as compared with the observations and Schecter function fits of Bouwens et al. (2015), at $\lambda = 1600\mathring{A}$. We use the escape fractions computed for each halo in calculating the total ionizing luminosity, which is then converted into a total magnitude which is used to compute the luminosity function. Due to the limited volume of our simulations, our results intersect only the faintest data point at $z = 6$, however the normalisation is consistent with the Schecter function fits at each redshift shown.

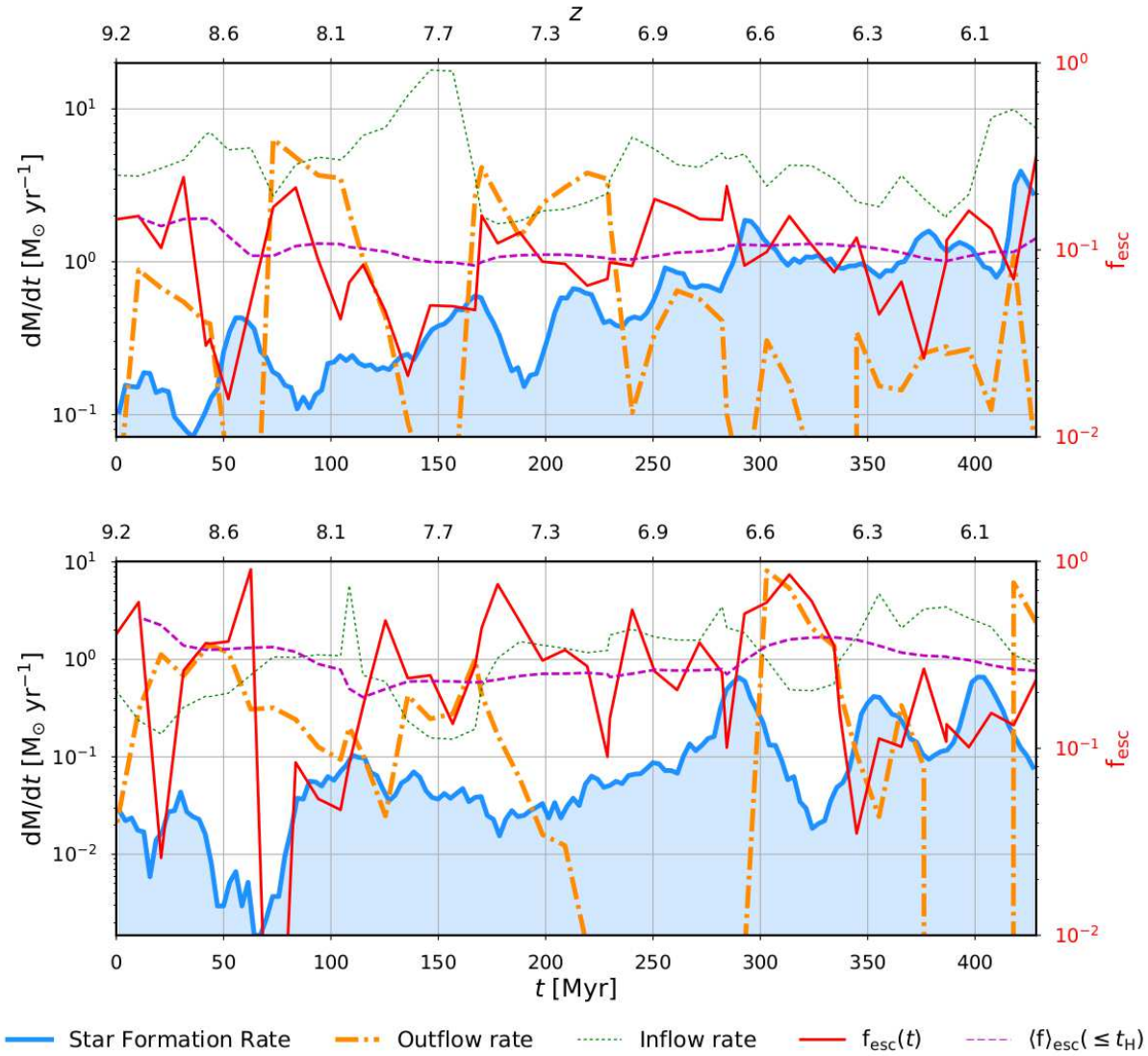

Figure 4.10: Time evolution of the star formation, outflow, and inflow rate combined with the instantaneous and photon-production weighted time-integrate escape fractions for two haloes with masses $M_{\rm vir} \simeq 6.29 \times 10^9\ M_\odot\ h^{-1}$ and $M_{\rm vir} \simeq 2.11 \times 10^9\ M_\odot\ h^{-1}$, respectively, from our binary population simulation. Episodes of star formation which trigger a peak in SNe activity produce strong outflows. On average, this also leads to a peak in $f_{\rm esc}$, which hints that the escape of ionizing photons is driven by strong outflows, although $\langle f_{\rm esc} \rangle (\leq t_{\rm H})$ remains roughly constant throughout.

feedback following the peak in star formation. While we have shown only two haloes here, we note that the same relationship between the star formation rate, outflow rate, and escape fraction is also observed in our single stellar population model. To demonstrate this, and further explore the relationship between the outflow rate and escape of ionizing photons, we plot in figure 4.11 the integrated number of photons escaping the virial sphere, as a function of final halo virial mass (left panel), and as a function of the integrated mass outflow rate from the central galaxy (right panel). The scatter shows the raw data points for each halo while the error bars show the mean and standard deviation, and $t_{\rm H}$ corresponds to the full age of the halo (which varies along the merger tree).

The two models show good agreement at low mass, with the binary population model predicting a larger on average number of ionizing photons escaping from high mass haloes, however both models are consistent within 1σ with only minor tension. The ratio of the total number of ionizing photons (binary to single) escaped into the IGM, integrated from $z \sim 9$ to $z \sim 6$, is only 1.73, suggesting that the inclusion of binaries do not significantly enhance the ionizing luminosity of haloes during reionization. In the right hand panel, we find a power-law fits the data well, with the single stellar population deviating marginally for high outflow rates (corresponding to high mass haloes). Furthermore, we find that the marginal reduction in stellar abundance, as shown in figure 4.8, has little to no impact on the total outflow rate, meaning SNe feedback is neither enhanced or diminished in either model (however the maximal outflow rate is still found to be larger in the single stellar population model).

The evidence of a power-law fit between the two quantities shown in the right hand panel strongly suggests that SNe driven winds regulate the escape of ionizing radiation. Remarkably, the efficiency of this process appears to remain constant over both halo mass and outflow rate, however drops marginally at high halo mass. Their simulations are part of the Feedback in Realistic Environment project (FIRE Hopkins et al., 2014), which were run using a pressure-entropy formulation of the smoothed particle hydrodynamics code GIZMO (Hopkins, 2015). Although they include a detailed treatment of feedback in their base cosmological simulations, the radiative transfer step is performed in post-processing, potentially neglecting any back-reaction of the radiation on the gas. They conclude that binary stars could boost the "effective" escape fraction (ratio of escaped ionizing photons to observed 1500 $\mathring{A}$ photons) by factors of $\sim 4 - 10$, however they note that important uncertainties remain in their work.

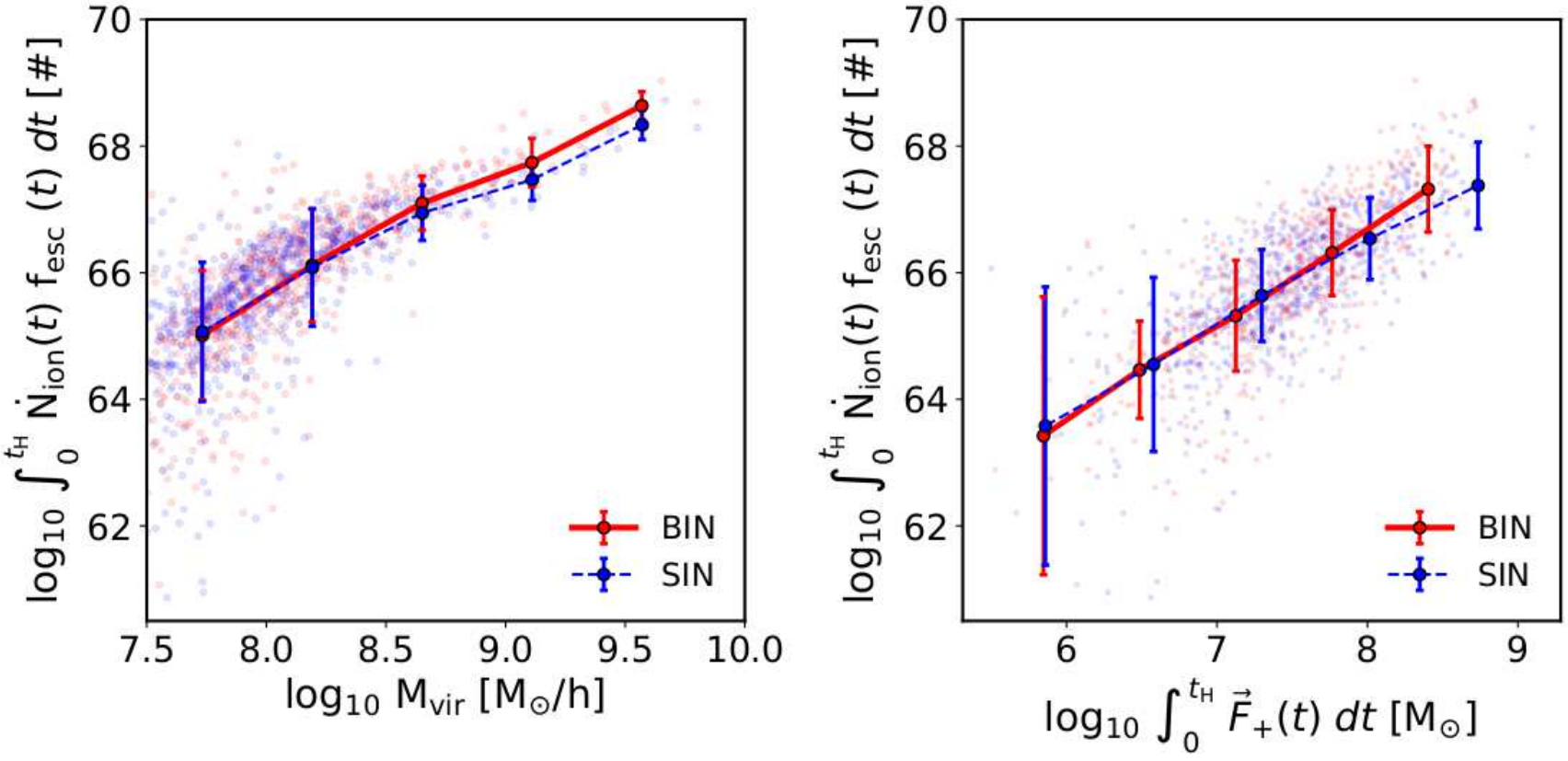

Figure 4.11: *Left panel:* Integrated number of ionizing photons escaping the virial sphere against halo virial mass. The points show the scatter in the data, while the error bars show the mean and 1σ standard deviation. The fit resembles a broken power law, with the knee at $M_{\rm vir} \sim 10^{8.5}\ M_\odot\ h^{-1}$. *Right panel:* The same as the left panel but plotted against the integrated outflow rate over the history of all haloes. The data follows a power law fit, suggesting that the efficiency of feedback-regulated escape of ionizing radiation remains roughly constant within the mass range shown.

4.5 Discussion

The escape fraction of galaxies during the EoR is one of the most crucial parameters in any model of reionization, however it remains not only poorly constrained but also poorly defined. Recently, Ma et al. (2016) showed using semi-numerical techniques that binary stars could in theory provide the 'missing photons' required for galaxies alone to drive cosmic reionization. Their study is based on post-processing RT of three galaxies, selected from a suite of cosmological zoom simulations where they use BPASS to model the stellar evolution of their galaxies.

In this work, we test the robustness of their conclusions by performing fully-coupled radiation hydrodynamical simulations of the EoR. In our simulations, the radiative transfer and hydrodynamical steps are fully coupled, therefore capturing any back-reaction on galaxies. Furthermore, due to our relatively large level of refinement, we are able to resolve individual galaxies in a cosmological context, hence better capturing the statistical deviation in $f_{\rm esc}$. The mass dependence of the escape fraction in our work are in good agreement with those of Kimm and Cen (2014); Kimm et al. (2017); Trebitsch et al. (2017), who used the same simulation code and similar methods for estimating $f_{\rm esc}$. However, we find very little or no redshift evolution, while Kimm and Cen (2014) find that the instantaneous $f_{\rm esc}$ gradually declines toward higher redshift for all halo masses. In their models, they developed and tested a new momentum-driven scheme for SNe feedback, which is shown to capture the Sedov blast wave solution, especially at low resolution. While the kinetic feedback scheme we use is known to perform better than traditional thermal dump methods, the lack of redshift evolution in our results could be due to the differences in SNe models chosen. On another detail of our method, Trebitsch et al. (2017) recently showed that for a method very similar to ours, there appears to be very little discrepancy with their more robust ray-tracing based method. This is largely due to the high central concentration of stars within their host halo, making the time-delay based method employed in this work a good approximation.

An interesting property of binary stars is that the shape of their UV spectrum hardens with population age. These harder photons have the potential to penetrate deeper into cold (neutral) pockets of gas, thus increasing the net photoheating rate. Furthermore, if these photons can raise the average gas temperature at the mean density, it has the potential to greater influence the growth of structure as compared with single stellar populations. Binary stars are able to raise the filtering mass (evaluated at the mean density of the Universe) throughout the EoR. While this effect is marginal, it leads to a subtle increase

in the effectiveness of radiative feedback on star forming galaxies, reducing their mean stellar abundance and UV luminosity. To keep the Universe ionized, any reduction in the density of stars, ρ_*, must therefore be met with a rise in the minimum escape fraction, as (Loeb and Furlanetto, 2013):

$$f_{\rm esc,min} \sim 1.6 \times 10^6 \left(\frac{\rho_*}{1\ M_\odot\ {\rm Mpc}^{-3}} \right)^{-1}. \tag{4.6}$$

While our results show that $f_{\rm esc}$ is indeed boosted due to binary stars, their increased suppression offsets much of this effect, leading to only a subtle enhancement (if any) in the number of ionizing photons, which escape to the IGM. The reason for this, however, is not due to a reduction in the number of SNe events, as the integrated number of escaping ionizing photons does not vary as a function of the integrated mass outflow rate. It should be noted that to truly confirm this result, larger boxes are required to better capture the total source population, as demonstrated by the lack of very bright objects in our UV luminosity functions, with respect to the observational constraints of Bouwens et al. (2015). We do, however, capture the faint-end of the UV luminosity function, for which our simulations predict a flattening out with respect to the predicted Schecter function fits. Observational challenges mean that the faint-end slope is poorly constrained, although we note that there is (albeit limited) evidence for such a flattening out at $z \sim 2-3$ (e.g Parsa et al., 2016).

On the tension between observations of the local Universe and the $f_{\rm esc} \sim 20\%$, required to match the Thompson scattering optical depth, we find that this arises due to the mass evolution of the halo population (increasing with time). Kimm et al. (2017) recently performed extremely high-resolution simulations, to probe the escape fraction of mini-haloes below our mass resolution, using detailed molecular chemistry to properly resolve the relevant physical cooling processes. They found that $f_{\rm esc}$ continues to increase towards lower masses than those shown in our results, and that the escape fraction is well approximated by a broken power law fit with:

$$\log f_{\rm esc} = \begin{cases} 0.17 - 0.2\ \log M_{\rm vir} & (M_{\rm vir} > 10^{8.5}) \\ 9.52 - 1.3\ \log M_{\rm vir} & (10^{7.75} < M_{\rm vir} \leq 10^{8.5}) \\ 1.0 - 0.2\ \log M_{\rm vir} & (M_{\rm vir} \leq 10^{7.75}). \end{cases}$$

Interestingly, their knee at $\log M_{\rm vir} = 10^{8.5}$ is in extremely good agreement with our photon-production weight time averaged escape fraction, as shown in figure 4.5. To properly convergence test our results, such a detailed treatment of molecular chemistry would

be required to properly model cooling processes in haloes below the atomic cooling limit (i.e our resolution limit). However, we do note that Kimm et al. (2017) use the same simulation code as us, only with a more detailed modelling of such haloes, therefore we do not expect significant deviation from their findings.

Provided the that the ability for SNe driven outflows to regulate $f_{\rm esc}$ does not diminish to higher halo masses, a linear extrapolation of our results predicts that a Milky Way sized halo will have $f_{\rm esc} \lesssim 1.5\%$. This is in good agreement with observations of our local Universe (e.g Leitet et al., 2011, 2013; Kuhlen and Faucher-Giguère, 2012; Finkelstein et al., 2012; Robertson et al., 2013, 2015), whilst simultaneously allowing for high enough escape fractions at early times to satisfy the requirements of SAMs (e.g Kuhlen and Faucher-Giguère, 2012; Finkelstein et al., 2012; Robertson et al., 2013, 2015).

4.6 Conclusions

We have presented the first RHD simulations which explore the impact of binary stars on the escape fraction of ionizing radiation. Assuming that $f_{\rm esc}$ is mass dependent only, as appears the case, then extrapolation of our results to the Milky Way dark matter halo mass leads to a prediction of $f_{\rm esc} \sim 1.55\%$ for the binary model, and $f_{\rm esc} \sim 1.20\%$ for the single stellar population model. A simple mass-dependent escape fraction can alleviate the tension between observations of the local Universe, which predict very low escape fractions, and simulations of the EoR which require $f_{\rm min,esc} \sim 20\%$.

Our findings are in strong agreement with Kimm and Cen (2014); Kimm et al. (2017); Trebitsch et al. (2017), who along with Razoumov and Sommer-Larsen (2006, 2007); Gnedin et al. (2008); Paardekooper et al. (2011); Kim et al. (2013) predict a declining $f_{\rm esc}$ with increasing halo mass. The reason for such a relationship is due to the strong dependence on SNe driven winds, which regulate $f_{\rm esc}$ by disrupting dense absorbers near the birth cloud. As the maximal density is expected to grow with mass of the host halo, increasingly powerful outflows are required to facilitate the escape of ionizing radiation.

The primary aim of this work, however, was to explore whether binary stars could provide the 'missing photons' required for galaxies to keep the Universe ionised. Recent semi-numerical studies by Ma et al. (2016) showed that, in theory, binary stars could provide the additional photons required to match observations. As they performed the RT step in post-processing, we conclude that they miss the small (but significant) additional suppression of star formation due to radiative feedback. Binary stars modify the shape of the SED of SSPs, leading to a hardening of their UV spectrum, which in turns leads to

additional heating during the early phases of reionization. As binary stars release the bulk of their photons, $\sim 60\%$ in their first 3 Myr of evolution compared to $\sim 40\%$ for single stellar populations, the initial growth of H II regions is also faster, which leads to a larger on average gas temperature.

We showed that this led to an increase in the filtering mass throughout reionization, which in turn leads to a small suppression in the growth of structures. This results in a (marginally) reduced stellar abundance on average, which acts to negate much of the boost in the number of ionizing photons escaping to the IGM and impact the faint-end UV luminosity function. We therefore conclude that while binary stars can provide additional photons for reionization, the predictions of Ma et al. (2016) overestimate their significance due to the semi-numerical nature of their calculations. In summary, we conclude that while binary stars can in theory provide more ionizing photons, galaxies become self-regulating over time hence limiting any potential boost in the escape fraction or number of ionizing photons which reach the IGM.

Acknowledgements

The authors contributed in the following way to this paper. DS undertook the calibration and completion of all simulations; the development of all analysis codes, the production of all figures, and the first draft of the paper. II supervised the project and helped with the interpretation of results and feedback on the first draft.

DS would like to thank Joakim Rosdahl for valuable technical feedback, Scott J. Clay and Benoît Fournier for useful discussions and feedback.

This work used the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The DiRAC system is funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grant ST/H008519/1, STFC DiRAC Operations grant ST/K003267/1, and Durham University. DiRAC is part of the National E-Infrastructure. We acknowledge PRACE for awarding us computational time under project PRACE4LOFAR grant 2014102339 and "Multi-scale simulations of Cosmic Reionization" grants 2014102281, 2015122822, and 2016153528 to resource Curie based in France at CEA and to resource Marenostrum based in Spain at BSC. This work was supported by the Science and Technology Facilities Council [grant numbers ST/F002858/1 and ST/I000976/1] and the Southeast Physics Network (SEPNet). Some of the analysis was done on the Apollo cluster at The University of Sussex.

We made extensive use of the PYMSES (Guillet et al., 2013) and PYNBODY (Pontzen et al., 2013) python modules when developing our own analysis pipeline.

Chapter 5

The Supersonic Baryon Drift Velocity

In this chapter, we present for the first time, a method for including the supersonic relative drift velocity, which exists between baryons and CDM at recombination, and its suppression in the growth of density models, in volumes larger than the coherence scale for the first time. To achieve this, we compute a scale dependent bias parameter, in sub-regions where the drift velocity is coherent, and convolve this into the underlying density fields of both dark matter and baryons as a sub-grid approximation. We find that while including the drift velocity predicted via linear theory does marginally suppress the halo mass function and global star formation rate, much of this effect is missed without proper modulation of the power spectrum, as done in our methodology. The work in this chapter is to be submitted for publication in Sullivan, D., Fialkov, A & Iliev, I. T (in prep.).

5.1 Introduction

As previously discussed in section 1.4, at the time of recombination baryons and photons decouple, and the sound speed in the baryonic fluid drops from relativistic, $\sim c/\sqrt{3}$, to the thermal velocities of hydrogen atoms, $\sim 2 \times 10^{-5} c$. This is less than the relative velocities of baryons and CDM predicted by linear perturbation theory, thus we infer that there existed a supersonic flow of baryons relative to the underlying potential wells generated by dark matter haloes (Tseliakhovich and Hirata, 2010). The mass scale of CDM perturbations, which grew to form the first collapsed objects at $z < 100$, crossed the horizon during the radiation dominated era when dark matter had already decoupled from the radiation plasma. Therefore, their growth was largely unaffected by the relative

streaming of baryons.

In the absence of weak structure formation shocks, the IGM is thought to be kinetically cold during the dark ages (prior to reheating by stars and galaxies), cooling adiabatically as $(1+z)^2$ and reaching 10 K by $z = 20$. However, as pointed out by O'Leary and McQuinn (2012), even 0.3 km s^{-1} flows would have been supersonic for gas at a temperature of 10 K. The evolution of density and velocity perturbations in the early Universe is generally described using linear perturbation theory (see section 1.3.1), which treats these quantities as small perturbations against the background, and neglects second order terms. However, as demonstrated by Tseliakhovich and Hirata (2010), the relative velocity between dark matter and baryons introduces an important second-order term in the Euler equations. At recombination, the gas was moving coherently on comoving scales $\lesssim 10$ Mpc, with a root-mean-square (rms) speed of $v_{bc} \sim 10^{-4}c = 30$ km s^{-1}, relative to dark matter. This translates to Mach numbers of $\mathcal{M}_{bc} \approx 1.7$ over the redshift range $15 \lesssim z \lesssim 150$, with a standard deviation of 0.7 (Tseliakhovich and Hirata, 2010; O'Leary and McQuinn, 2012). As such, supersonic flows could source shocks and entropy generation, which potentially has significant implications for the growth of perturbations following recombination. Furthermore, it could lead to inhomogeneous heating in the early Universe, and coupling to the 21cm signal as a result (e.g McQuinn and O'Leary, 2012).

This could have important consequences on whether baryons are able to collapse into the first haloes and form stars, as the relative velocity acts as an effective sound speed raising the filtering mass at early times (see figure 1.4, courtesy of Tseliakhovich et al., 2011). The growth of these dark matter haloes is therefore inhibited, as baryons are advected down-wind of the halo forming Mach-cones, which suppresses the halo mass function at the low mass end; an effect first shown in simulations by O'Leary and McQuinn (2012). Therefore, this could have a significant impact on the formation of the first Population III stars in the Universe, and hence reionization (e.g Maio et al., 2011; Stacy et al., 2011; Greif et al., 2011; O'Leary and McQuinn, 2012; Fialkov et al., 2012).

On scales where the relative velocity is coherent, the Euler equations can be linearised, to properly solve for the growth of perturbations. This was first done by Tseliakhovich and Hirata (2010), and subsequently included in simulations with small volumes below the coherence scale (e.g Maio et al., 2011; Stacy et al., 2011; Greif et al., 2011; Naoz et al., 2012; O'Leary and McQuinn, 2012; Naoz et al., 2013). On these scales, the drift velocity is typically included as a coherent wind across the entire box, allowing one to linearise the Euler equations. To fully capture this effect in much larger, cosmologically

representative volumes, would require an accurate non-linear treatment of dark matter, baryons, and radiation starting at $z_{\rm rec} \sim 1000$. This is a feat which is currently not possible, as the amplitude of density fluctuations at these times become overwhelmed with shot noise (coupled with the lack of accurate treatment of the radiation dominated era in such codes).

In this chapter, we introduce a new method, based on the work of O'Leary and McQuinn (2012), to include the relative velocity in cosmological initial conditions with volumes many times larger than the coherence scale, for the first time. The pre-reionization Universe is observable via the redshifted 21 cm line in absorption against the CMB, which is tightly coupled to the thermal history of the IGM and the star formation rate during this epoch (e.g Madau et al., 1997; Furlanetto et al., 2006; Furlanetto, 2006; McQuinn and O'Leary, 2012), therefore the ability to simulate this effect in volumes comparable to the survey size, is of great importance to understanding future observations (i.e SKA). The key focus of this chapter, is to introduce a method for modulating the power spectrum of fluctuations, in both dark matter and baryons, due to the relative drift velocity in several-hundred Mpc volumes comoving, where the BAOs (which source the relative velocity and have characteristic scales of ~ 100 Mpc) are captured.

In section 5.2, we introduce our methodology and demonstrate our ability to generate high-resolution, zoom initial conditions, capable of resolving the relevant scales. Specifically, in section 5.2.6, we define our scale-and drift velocity-dependent *bias parameter*, $b(k, v_{bc})$, and describe the process of convolving this bias into our initial conditions. In section 5.3, we present preliminary results from the first large-scale simulations with this effect included, before discussing our findings and conclusions in section 5.4.

5.2 Method

5.2.1 Theory

In this section, we briefly review the relevant theoretical background, however direct the reader to Fialkov (2014) for a more comprehensive review. As a reminder, we adopt the notation $\mathcal{P}(k)$ for the dimensionless power spectrum, usually denoted Δ^2.

In the non-linear regime, the full non-linear equations to solve are:

$$\begin{aligned}
&\frac{\partial \delta_c}{\partial t} + a^{-1}\mathbf{v}_c \cdot \nabla \delta_c = -a^{-1}(1+\delta_c)\nabla \cdot \mathbf{v_c}, \\
&\frac{\partial \delta_b}{\partial t} + a^{-1}\mathbf{v}_b \cdot \nabla \delta_b = -a^{-1}(1+\delta_b)\nabla \cdot \mathbf{v_b}, \\
&\frac{\partial \mathbf{v}_c}{\partial t} + a^{-1}(\mathbf{v}_c \cdot \nabla)\mathbf{v}_c = -\frac{\nabla \Phi}{a} - H\mathbf{v}_c, \\
&\frac{\partial \mathbf{v}_b}{\partial t} + a^{-1}(\mathbf{v}_b \cdot \nabla)\mathbf{v}_b = -\frac{\nabla \Phi}{a} - H\mathbf{v}_b - a^{-1}c_s^2 \nabla \delta_b, \\
&a^{-2}\nabla^2 \Phi = 4\pi G \bar{\rho}_m \delta_m,
\end{aligned} \tag{5.1}$$

where $\mathbf{v_b}$ and $\mathbf{v_c}$ are the velocities of baryon and dark matter respectively, a is the scale factor, $H \equiv \dot{a}/a$ is the Hubble constant, and

$$c_s^2 = \frac{k_B T_K}{\mu}\left(1 - \frac{1}{3}\frac{\partial \log T_K}{\partial \log a}\right), \tag{5.2}$$

is the baryonic sound speed (which in general is non-uniform; Naoz and Barkana, 2005), with k_B the Boltzmann constant, T_K the kinetic gas temperature, and μ the mean molecular weight. In the regions where the velocities are mostly coherent, we can split the velocities into $\mathbf{v_b} = \mathbf{u_b}$ and $\mathbf{v_c} = \mathbf{v_{bc}} + \mathbf{u_c}$. On scales where v_{bc} is coherent, the linear perturbations in density (δ_b and δ_c) and velocities ($\mathbf{u_b}$ and $\mathbf{u_c}$) evolve according to the following set of equations in Fourier space (after we subtract the coherent mode $\mathbf{v_{bc}}$, and define the comoving velocity divergence, $\theta_i = a^{-1}\nabla \mathbf{u_i}$):

$$\begin{aligned}
&\frac{\partial \delta_c}{\partial t} = \frac{i}{a}\mathbf{v}_{\rm bc} \cdot \mathbf{k}\delta_c - \theta_c, \\
&\frac{\partial \delta_b}{\partial t} = -\theta_b, \\
&\frac{\partial \theta_c}{\partial t} = \frac{i}{a}\mathbf{v}_{\rm bc} \cdot \mathbf{k}\theta_c - \frac{3H^2}{2}(\Omega_c \delta_c + \Omega_b \delta_b) - 2H\theta_c, \\
&\frac{\partial \theta_b}{\partial t} = -\frac{3H^2}{2}(\Omega_c \delta_c + \Omega_b \delta_b) - 2H\theta_b + \frac{c_s^2 k^2}{a^2}\delta_b,
\end{aligned} \tag{5.3}$$

These are the modified equations, linear in terms of the real perturbations, properly describing the growth of structure when the relative velocities of baryons and dark matter are taken into account. In these equations, $\mathbf{v_{bc}}$ redshifts away as $(1+z)$ as it is a vector perturbation.

In the cold dark matter frame, these equations of motion become (see also the appendix of O'Leary and McQuinn, 2012):

$$\begin{aligned}
\frac{\partial \delta_c}{\partial t} &= -\theta_c, \\
\frac{\partial \delta_b}{\partial t} &= -\frac{i}{a}\mathbf{v}_{\rm bc}\cdot\mathbf{k}\delta_c - \theta_b, \\
\frac{\partial \theta_c}{\partial t} &= -\frac{3H^2}{2}(\Omega_c\delta_c + \Omega_b\delta_b) - 2H\theta_c, \\
\frac{\partial \theta_b}{\partial t} &= -\frac{i}{a}\mathbf{v}_{\rm bc}\cdot\mathbf{k}\theta_c - \frac{3H^2}{2}(\Omega_c\delta_c + \Omega_b\delta_b) - 2H\theta_b + \frac{c_s^2k^2}{a^2}\delta_b.
\end{aligned} \tag{5.4}$$

We use the CICSASS code (O'Leary and McQuinn, 2012) to solve the above set of linear equations for fixed v_{bc} at recombination (hereafter $v_{bc,\rm rec}$), and compute the resulting density power spectra for both baryons and CDM. We then define our scale- and v_{bc}-dependent bias parameter, $b(k, v_{bc})$, which is simply the ratio of the modulated power spectrum with the linear theory predicted case ($v_{bc} = 0$):

$$b(k, v_{bc}) = P(k, v_{bc})/P(k, v_{bc} = 0). \tag{5.5}$$

This bias is interpolated onto 3D grids, corresponding to each level in our zoom hierarchy, and convolved into the initial δ_c and δ_b fields. Note, that the linear theory v_{vc} comes naturally from subtracting the baryon and CDM velocity fields produced by MUSIC, as it originates from the BAOs.

The importance of gas pressure on the growth of density modes was stressed by Naoz and Barkana (2005, 2007), which requires an additional equation for the temperature fluctuations required to solve for c_s:

$$\frac{d\delta_T}{dt} = \frac{2}{3}\frac{d\delta_b}{dt} + \frac{x_e(t)}{t_\gamma}a^{-4}\left[\delta_\gamma\left(\frac{\bar{T}_\gamma}{\bar{T}} - 1\right) + \frac{\bar{T}_\gamma}{\bar{T}}(\delta_{T_\gamma} - \delta_T)\right] \tag{5.6}$$

where $x_e(t)$ is the electron fraction at time t, δ_γ is the photon density fluctuation,

$$t_\gamma^{-1} = \frac{8}{3}\bar{\rho}_\gamma^0 \times \frac{\sigma_T c}{m_e} = 8.55 \times 10^{13}\ yr^{-1},$$

and T_γ and δ_{T_γ} are the mean photon temperature and its dimensionless fluctuation, respectively. While the inclusion of this equation is important to properly account for the full thermal history of the gas, at the time of writing this has not been implemented in our initial conditions, therefore this is left for future work.

5.2.2 Cloud in Cell Smoothing

In order to estimate the CDM density field, we use Cloud in Cell (CIC) smoothing to assign particles onto each point of a regular cartesian grid. For an N-body system of N_p DM particles, the particle number density is given by:

$$n_0(\mathbf{x}) = \sum_{i=1}^{N_p} \delta^D(\mathbf{x} - \mathbf{x}_i), \tag{5.7}$$

where $\mathbf{x}_i$ is the position of the i-th particle, and δ^D is the Dirac delta function. We sample the continuous number density, $n^s(\mathbf{x})$, onto regular grids of size $N_{\rm grid}$ in three dimensions as:

$$n^s(\mathbf{x}_p) = \int_V d^3x' n_0(\mathbf{x}') W(\mathbf{x}_p - \mathbf{x}'), \tag{5.8}$$

where $W(\mathbf{x})$ is the window function which quantifies how particle number density is assigned onto grid points separated by $\mathbf{x}$. Therefore the sampled density contrast, $\delta^s(\mathbf{x}) = n^s(\mathbf{x})/\bar{n} - 1$, is given by the real space linear convolution:

$$\delta^s(\mathbf{x}) = [\delta * W](\mathbf{x}), \tag{5.9}$$

where δ is the real continuous density field, which in Fourier space simplifies to the product:

$$\delta^s(\mathbf{k}) = \delta(\mathbf{k}) W(\mathbf{k}). \tag{5.10}$$

To define the window function, we must first define the cloud shape function (or point spread function), $S(\mathbf{x}')$, for which a particle with unit mass gives its mass density, where x' measures the distance from the centre of the particle. In the one-dimensional case, the fraction of mass assigned from a particle of shape S, at position x, to the grid point p, at x_p, is then given by the overlap of the cloud with cell p (Hockney and Eastwood, 1988):

$$W(x - x_p) = W_p(x) = \frac{1}{H} \int_{x_p - H/2}^{x_p + H/2} S(x' - x) dx', \tag{5.11}$$

where $H = L/N_{\rm grid}$ is the size of the regular grid element. By defining the top-hat function as:

$$\Pi(x) = \begin{cases} 1 & \text{if } |x| < 1/2 \\ 1/2 & \text{if } |x| = 1/2 \\ 0 & \text{otherwise} \end{cases},$$

and using equation 5.11 we can define the relationship between the cloud shape S and assignment function W:

$$W(x) = \frac{1}{H}\int \Pi\left(\frac{x'}{H} * S(x)\right). \tag{5.12}$$

The simplest possible assignment scheme is Nearest Grid Point (NGP), for which the one-dimensional window function is proportional to the top-hat function:

$$W_{\mathrm{NGP}}(x) = \frac{1}{H}\Pi\left(\frac{x}{H}\right), \tag{5.13}$$

which, in Fourier space is simply:

$$W_{\mathrm{NGP}}(k) = \mathrm{sinc}\left(\frac{\pi k}{2k_N}\right), \tag{5.14}$$

where $k_N = \pi/H$ is the Nyquist frequency. The CIC assignment scheme is the first order distribution scheme, which uniformly distributes particles with a top-hat spreading function. As such, the spreading function is defined as:

$$S_{\mathrm{CIC}}(x) = \frac{1}{H}\Pi\left(\frac{x}{H}\right), \tag{5.15}$$

and the window function is given by the convolution (in Fourier space):

$$W_{\mathrm{CIC}}(k) = W_{\mathrm{NGP}}(k)^2 = \mathrm{sinc}^2\left(\frac{\pi k}{2k_N}\right). \tag{5.16}$$

As the window function is well defined, the real density field is trivially recovered by deconvolution of W_{CIC} with the sampled density field, δ^s following equation 5.10, to recover small-scale (high k) power lost in the smoothing process. In three dimensions, the window function is simply the product of three one-dimensional window functions, i.e:

$$W(\mathbf{x}) = W(x, y, z) = W(x)W(y)W(z), \tag{5.17}$$

which in Fourier space leads to the general case:

$$W(\mathbf{k}) = \left[\mathrm{sinc}\left(\frac{\pi k_x}{2k_N}\right)\mathrm{sinc}\left(\frac{\pi k_y}{2k_N}\right)\mathrm{sinc}\left(\frac{\pi k_z}{2k_N}\right)\right]^p, \tag{5.18}$$

where $p = 1, 2$ for NGP and CIC respectively.

5.2.3 Initial Conditions

We utilise the MUSIC code (Hahn and Abel, 2013), which is based on the work of Bertschinger (2001), to generate high-resolution *zoom* initial conditions. These are capable of resolving the scales where the relative drift velocity modulates the power spectrum of density fluctuations, on scales $k \gtrsim 10$ Mpc^{-1} h (Tseliakhovich and Hirata, 2010; O'Leary and McQuinn, 2012), and in volumes of several-hundred Mpc comoving. To develop and test our method, we construct density and velocity fields in a 100 Mpc h^{-1} comoving volume at $z = 200$, initially assuming $v_{bc} = 0$, using transfer functions from CAMB (Lewis et al., 2000). The base grid level is set to $\ell = 9$, corresponding to 512^3 particles and grid cells, and we refine up to $\ell = 14$ giving an effective resolution of $16,384^3$ in the high resolution region. We use second-order Lagrangian perturbation theory (2LPT), as first-order LPT is known to underestimate the level of small-scale power with respect to 2LPT (see e.g Hahn and Abel, 2013; L'Huillier et al., 2014).

To select our zoom region, we perform an N-body only uni-grid simulation and select the largest halo at $z = 10$ for test purposes. The region of refinement is then determined by computing the Lagrangian volume of all particles form the halo, from the initial conditions at $z = 200$ to $z = 10$. We show the complete geometry in figure 5.1, focused on the centre of the box where the refinement takes place. The final region need not take on any special geometry, therefore we use a spheroid with radius $r \sim 3$ Mpc h^{-1} comoving, which bounds the halo of interest. A small buffer region is placed around each level of refinement, to prevent discontinuities developing at the refinement region boundaries.

This allows us to achieve a formal dark matter particle resolution of $1.43 \times 10^4 M_\odot\ h^{-1}$ in our 100 Mpc h^{-1} volume, able to resolve mini-haloes with masses $\sim 1 \times 10^6\ M_\odot\ h^{-1}$ with 70 particles. Most importantly however, is our ability to reach an effective cell size of 6.10352 kpc h^{-1} in our initial conditions[1], satisfying the requirement that we capture scales $k \gtrsim 10$ Mpc^{-1} h where the drift velocity modulates the power spectrum, whilst simultaneously including large-scale modes.

5.2.4 Power Spectrum of Density and Velocity Fields

We compute the density of dark matter by using a CIC smoothing scheme, where the point spread function is deconvolved from the 3D density field as described in the previous section, with the Zel'dovich approximation which relates the velocity field with the initial

[1]Note, that in our AMR simulations we refine the grid to much smaller spatial scales, able to resolve mini-haloes.

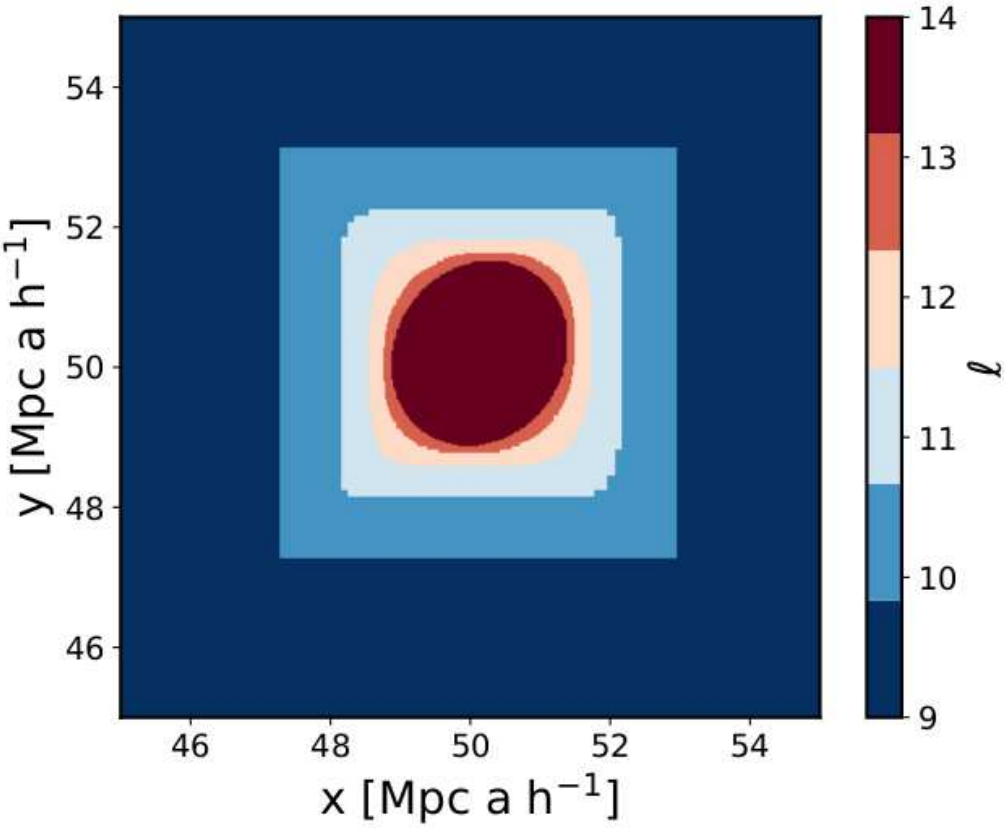

Figure 5.1: Geometry of our initial conditions, showing the high-resolution region at the centre of the box. The color-bar denotes the level in the AMR hierarchy, ℓ, where the effective number of particles and grid cells is given by $(2^{\ell})^3$. As shown, the geometry is not restricted in shape, where we have adopted an ellipsoidal region, which bounds all particles which form the largest halo at $z = 10$.

displacement field, as:

$$\mathbf{x} \equiv \mathbf{q} + \mathbf{\Psi} = \mathbf{q} - \nabla\phi, \tag{5.19}$$

where $\mathbf{\Psi}$ is the linearised displacement from Lagrangian position $\mathbf{q}$, and is related to the Eulerian linear perturbation theory overdensity, δ, by $\nabla \cdot \mathbf{\Psi} = -\delta$. The relationship to the velocity field comes from taking the time derivative:

$$\mathbf{v} = -a\nabla\dot{\phi}, \tag{5.20}$$

where $\nabla^2\phi = \delta$. In figure 5.2, we show slices through the baryon (left) and CDM (right) over density field, $\delta_{b,c}$ respectively.

In the linear regime, baryons trace dark matter as shown. We compute the density power spectrum in Fourier space using equation 1.22 and radially averaging over k modes, which we plot in figure 5.3 where the solid lines show the measured power spectra for the full box (left panel, at grid level $\ell = 9$) and the high resolution region (right panel, at grid level $\ell = 14$). The dotted lines show the theoretical power spectra using the transfer functions, $T(k)$, produced using the CAMB code and using equation 1.23 for baryons (orange) and CDM (blue). The dark blue dashed lines show the CDM power spectrum calculated without deconvolving the CIC window function, $W_{\rm CIC}(\mathbf{k})$. In the left hand panel, while the deconvolution does recover some small-scale power, the power spectrum is still suppressed at high-k due to the limited resolution in our CIC interpolated density field. Although CIC is known to perform better than NGP, it can still lead to discreteness effects as it still relies on NGP to assign particles to their nearest grid cells. This could be further improved by using a more sophisticated smoothing kernel, such as SPH, as cells receive a mass contribution from their nearest particles regardless of distance.

Velocity fields are constructed using species specific velocity transfer functions, from which the velocity power spectra are calculated using equation 6 of Iliev et al. (2007b), i.e:

$$P_v(k,t) = \frac{a^2(t)}{k^2}\left[\frac{\dot{D}_+(t)}{D_+(t)}\right]^2 P_\delta(k,t), \tag{5.21}$$

where $P_\delta(k,t)$ is the power spectrum of density fluctuations at time t. We show the theoretical velocity power spectrum in units of km s^{-1}, defined as:

$$\mathcal{P}_v(k,t) = \frac{k^3}{2\pi^2}P_v(k,t), \tag{5.22}$$

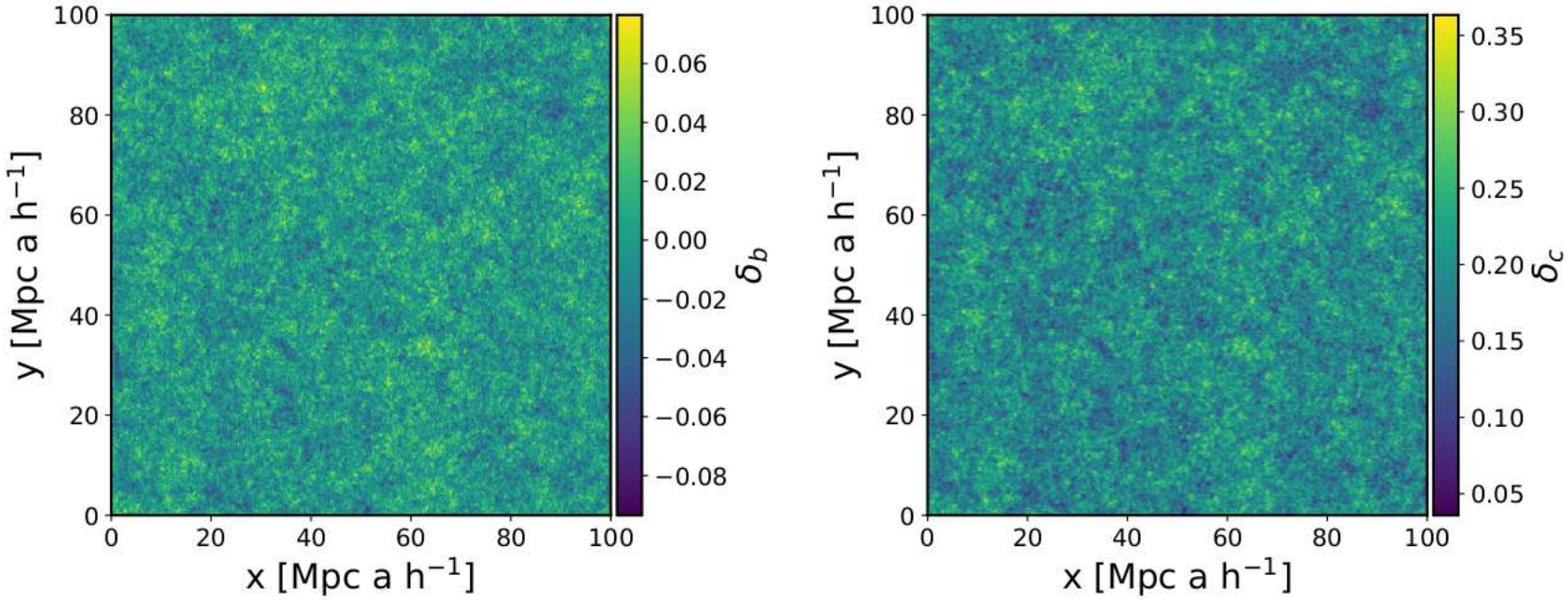

Figure 5.2: Slices through the baryon and CDM overdensity fields, δ_c and δ_b respectively, at $z = 200$. We compute the CDM overdensity by CIC smoothing N-body particles onto regular Cartesian grids as described in section 5.2.2. At this redshift, both fields are highly homogeneous and linear, as $\delta < 1$, and baryons roughly trace dark matter.

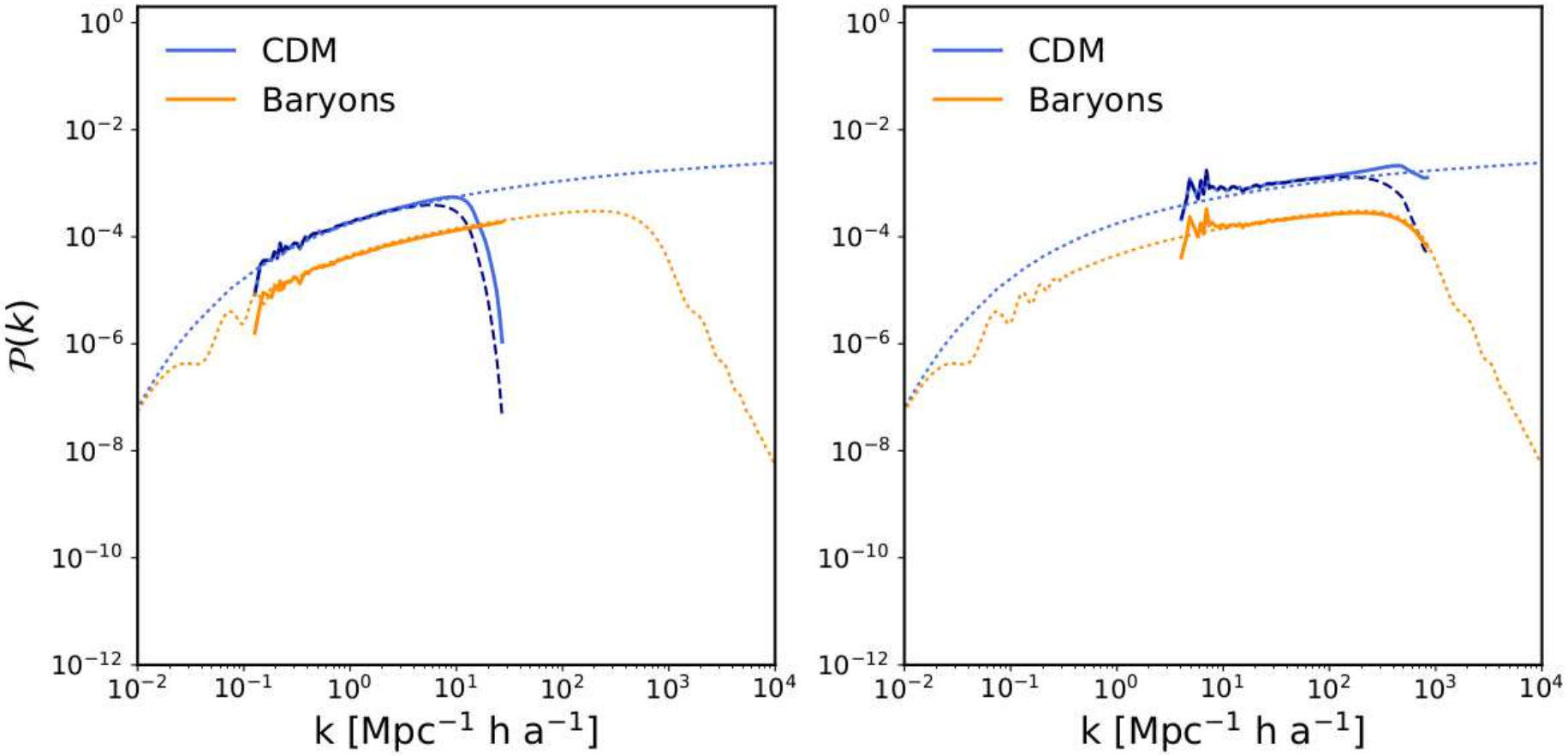

Figure 5.3: Dimensionless power spectra of density fluctuations for baryons (solid orange) and CDM (solid blue) in our initial conditions assuming $v_{bc} = 0$, compared with the theoretical prediction from CAMB (dotted lines), for the full box at $\ell = 9$ (left panel) and the high-resolution region at $\ell = 14$ (right panel). We also show the power power spectrum of the CDM field without deconvolution of the CIC window function, $W_{\rm CIC}$ (dashed navy line). The cut-off in baryon power at high-k is due to Silk damping.

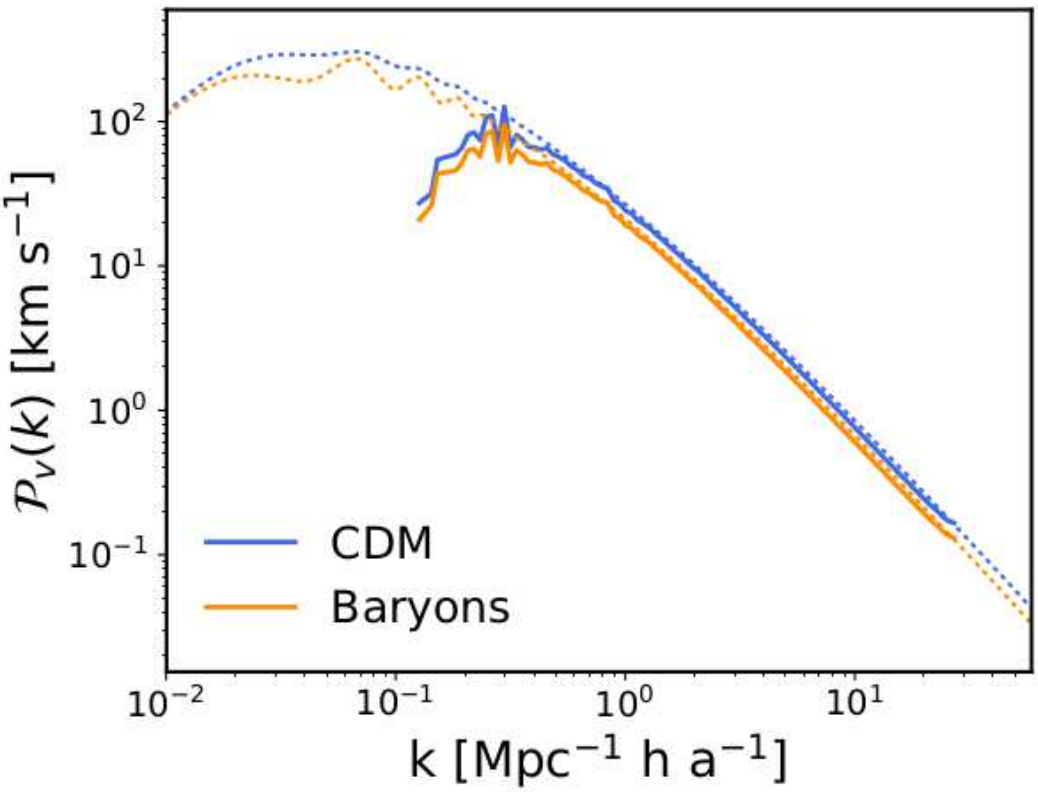

Figure 5.4: Velocity power spectra in units of km s^{-1} for the low-resolution region for baryons (solid orange) and CDM (solid blue), where the predicted power spectra (dotted lines) are calculated using equation 5.21. As the velocity fields of CDM and baryons are continuous fields in our initial conditions, we find good agreement with the theoretical spectra down to the smallest scales resolved. As discussed in the text, velocity perturbations are sourced by gravitational fluctuations in the density field at larger scales, therefore due to our limited volume, we find a loss of velocity power at large-scales. It should be noted, that in order to capture the peaks in the v_{bc} field, one must fully capture the BAO scales.

along with the measured velocity power spectra from our initial conditions in figure 5.4. The velocity field is trivially related to the density field via the Fourier transform of the Continuity equation:

$$\mathbf{u_k} = -i\frac{aHf(\Omega)}{k}\delta_{\mathbf{k}}\hat{\mathbf{k}}, \tag{5.23}$$

where $f(\Omega) = (a/D)(dD/da) \approx \Omega^{0.6}$ to a very good approximation (Loeb and Furlanetto, 2013). Therefore, peculiar velocity perturbations grow proportionally to density fluctuations, as $\mathbf{u_k} \propto \delta_{\mathbf{k}}/k$. The loss of power on large scales is due to Windowing, as the limited volume of our simulation box lacks large-scale modes.

The velocity power spectra of our initial conditions diverges from the theoretical predictions close to the BAO scale, where the drift velocity is expected to be strongest. For the purpose of developing and testing our methodology, we restrict ourselves to this volume, as the computational resources required to resolve these scales in larger boxes are not readily available. Therefore, all results presented here are likely lower limits.

5.2.5 Drift Velocity

The linear theory predicted drift velocity arises naturally in our initial conditions, due to the differing velocity transfer functions from which we construct the initial velocity fields. However, as previously discussed, this only constructs the initial v_{bc} field while missing its modulation of the growth of modes. The root-mean-square of the $|v_{bc,\mathrm{rec}}|$ field, i.e v_{bc} at recombination, in our initial conditions is measured to be 19.06 km s^{-1}, compared with the predicted 30 km s^{-1} by Tseliakhovich and Hirata (2010). The reason for this, as discussed in the previous section, is that we do not fully capture the BAOs which source such large v_{bc}. However, at recombination, this still leads to a drift with Mach number $\mathcal{M}_{bc} \approx 3$.

In figure 5.5, we show a slice through the $|v_{bc}| = |v_b - v_c|$ field at $z = 200$, for our full volume. As expected, v_{bc} appears largely coherent on scales $\lesssim 10$ Mpc $a\ h^{-1}$. The peak drift velocity, $|v_{bc,\mathrm{max}}| \simeq 8$ km s^{-1} at $z = 200$, and reaches zero in some regions. The impact on the growth of structure is therefore expected to largely dependent on environment, requiring large volumes to properly sample the statistical variation.

To properly calculate the coherence scale in our initial conditions, we compute the power spectrum of this field as shown in figure 5.6 in units of km s^{-1}. This was first calculated by Tseliakhovich and Hirata (2010), who showed that power begins to rapidly decline on scales $k \gtrsim 0.5$ Mpc^{-1} comoving as v_{bc} becomes coherent. Our findings show good agreement with theirs, however we once again note that larger boxes are required to

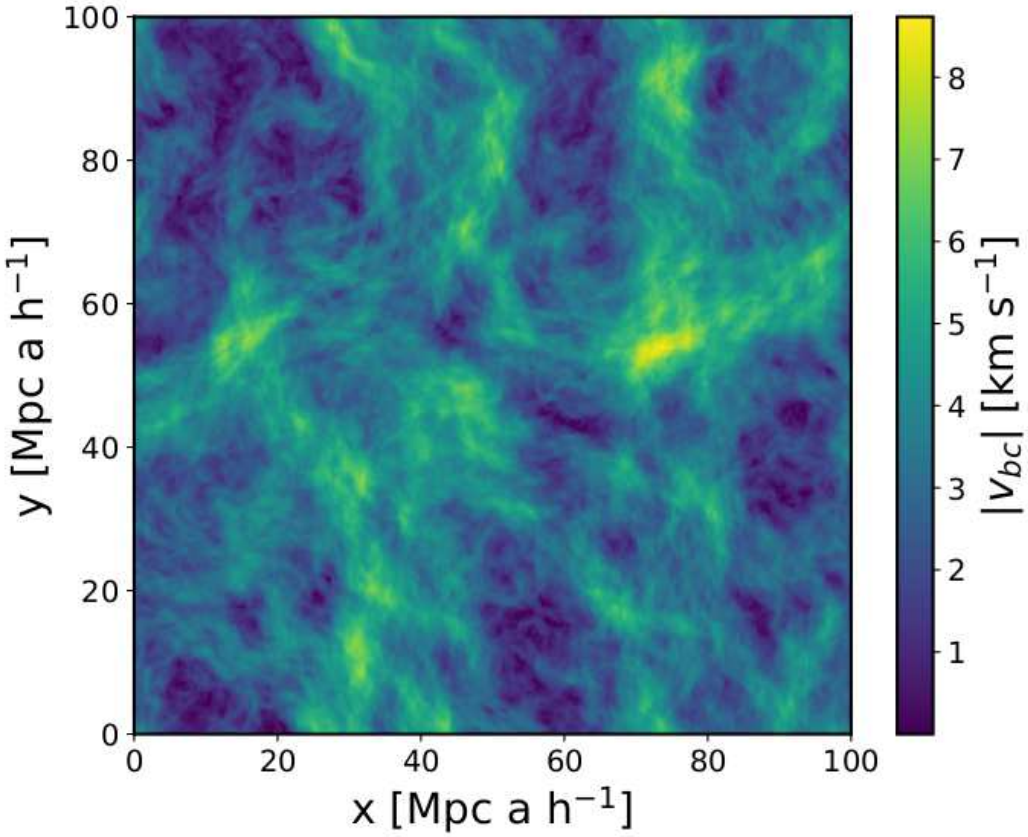

Figure 5.5: Slice through the $|v_{bc}| = |v_b - v_c|$ field at $z = 200$, of one cell thickness (6.10352 kpc h^{-1}). The drift velocity arises naturally when initialising the baryonic and CDM velocity fields with their own transfer functions.

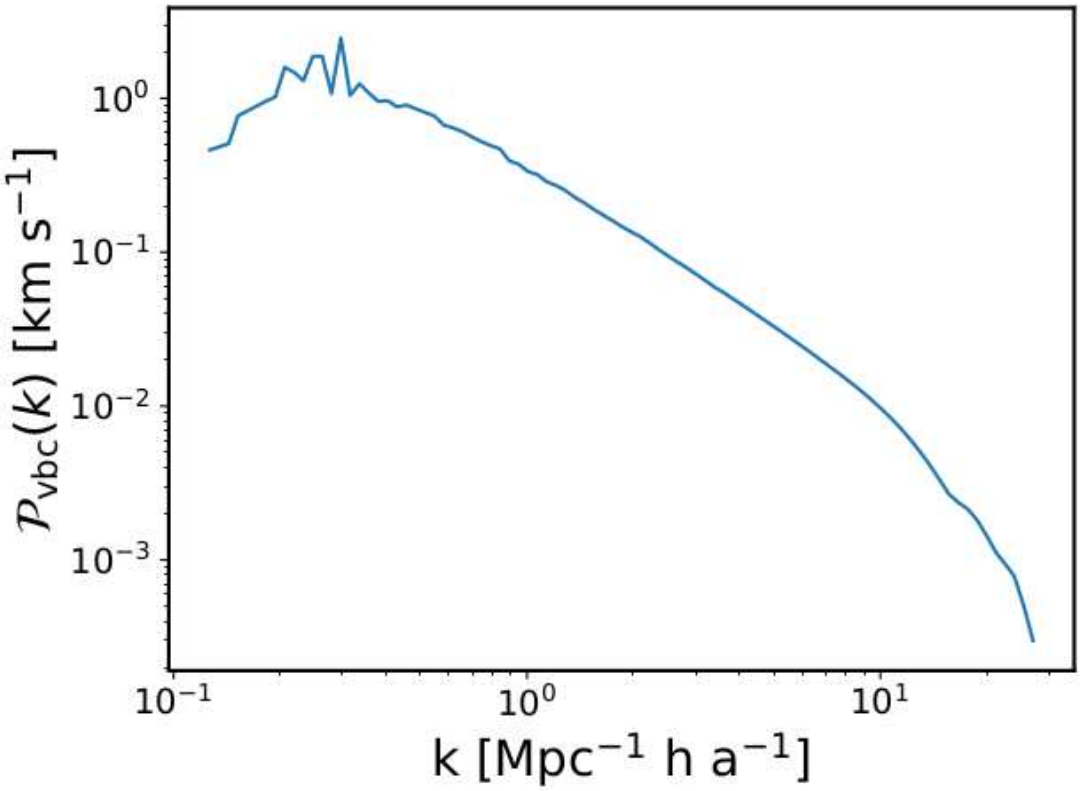

Figure 5.6: Power spectrum of the v_{bc} field, in units of km s^{-1}, calculated using the continuous baryon and CDM velocity fields from our initial conditions. Power begins to decline at $k \gtrsim 0.5$ Mpc^{-1} in good agreement with Tseliakhovich and Hirata (2010).

fully sample the large-scale modes.

5.2.6 Scale Dependent Bias

In this section, we discuss our methodology for computing the bias parameter $b(k, v_{bc})$. We split our initial conditions into cubic subregions each with size $\Delta x = 0.5$ Mpc a h^{-1} on-a-side, in which we use the system of equations listed in equation 5.4 with $v_{bc} = 0$ and the mean v_{bc} measured in each cubic sub-region. The bias is then calculated using equation 5.5, and interpolated onto a 3D grid for each $\mathbf{k}$ mode. This is done for both dark matter and baryons, after which the bias is convolved into the underlying matter density field in Fourier space as:[2]

$$\delta_m^b(\mathbf{k}) = \delta_m(\mathbf{k})\sqrt{b(\mathbf{k}, v_{bc})}, \tag{5.24}$$

where $\delta_m^b(\mathbf{k})$ is the *biased* overdensity field whose power spectrum has been modified. The biased velocity fields are also calculated using equation 5.23 and measuring the difference between the biased and unbiased fields and updating the velocity fields accordingly. In figure 5.7, we show $b(k, v_{bc})$ for two test regions with mean $v_{bc,\mathrm{rec}} = 11.56$ and 26.02 km s^{-1} (solid and dot-dashed lines, respectively), for baryons (orange) and CDM (blue) across a range of scales spanning all regions in our zoom hierarchy. The majority of suppression occurs on scales $k > 10^2$ Mpc^{-1} h a^{-1} for both baryons and CDM, which are only resolved in our high-resolution region. As expected, the magnitude of this suppression is proportional to the magnitude of the drift velocity field at recombination, with the scale at which suppression is strongest also shifting to lower-k for higher v_{bc}.

To prevent discreteness effects at boundaries, we pad each cubic sub-region with 8 cells when performing the above convolution. In the bottom panel of figure 5.8, we show the resulting biased power spectra (solid) of CDM and baryons compared to the CAMB linear theory power spectra (dotted) in our high-resolution region. Note, that while the rms of the entire region is measured to be $v_{bc,\mathrm{rec}} = 19.06$ km s^{-1} at recombination, the modulation of the power spectrum is done independently for each cubic sub-region. In the top panel, we also show the globally averaged bias of both species calculated by taking the ratio of the biased and unbiased power spectra shown in the lower panel. Both the globally averaged bias and biased power spectra show no evidence of discreteness effects at small-scales, and upon inspection the mean density in the high-resolution is unchanged to four decimal places at 3.45289×10^{-21} kg m^{-3} and 3.45283×10^{-21} kg m^{-3} for the

[2] We take the square root of the bias parameter as the power spectrum is defined as $\langle|\delta_{\mathbf{k}}|^2\rangle$.

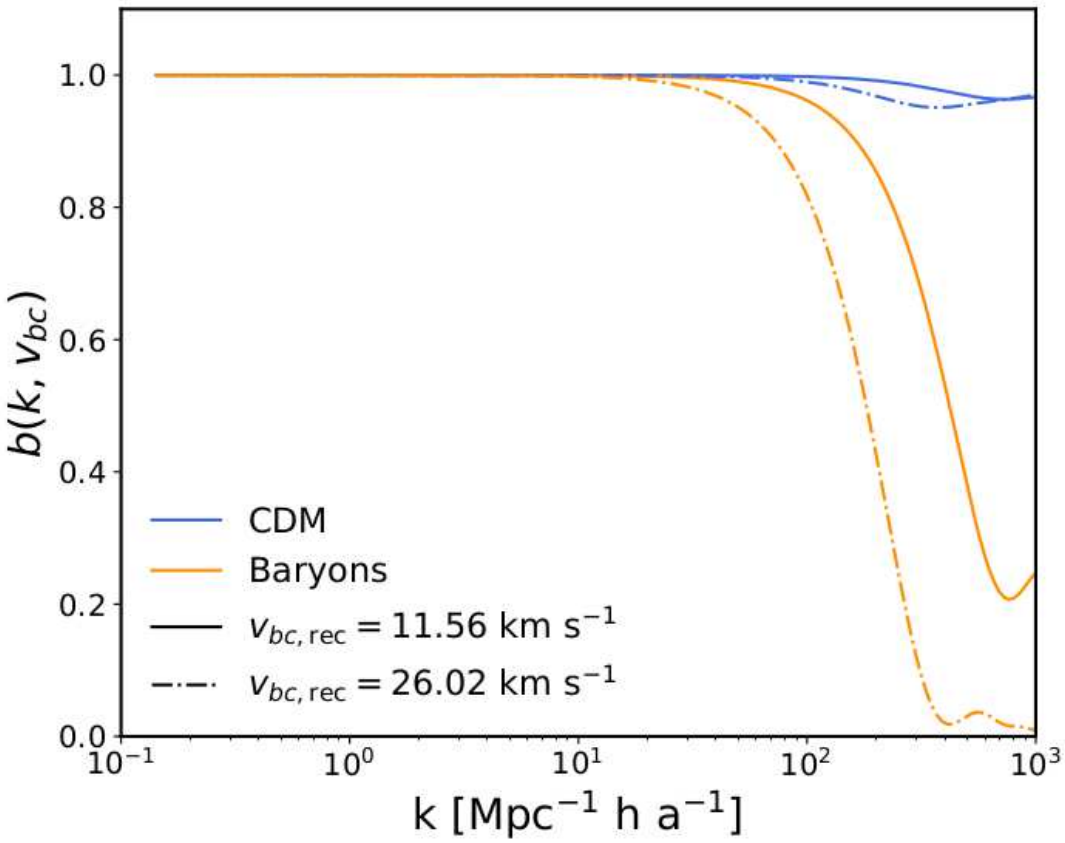

Figure 5.7: Scale and v_{bc} dependent bias, $b(k, v_{bc})$, for two different values of the drift velocity measured at recombination for each species individually. The degree by which power is suppressed is proportional to the magnitude of the drift velocity, with power on scales $k > 10^2$ Mpc^{-1} h a^{-1} most suppressed in both species.

unbiased and biased case respectively. While the modulation of the CDM power spectrum is marginal, the effect is captured when we recompute the density field following the Zel'dovich approximation.

5.2.7 Cosmological Simulations

We assume a geometrically flat ΛCDM Universe, consistent with a WMAP-7 cosmology (Komatsu et al., 2011) with cosmological parameters: $\Omega_{\rm M} = 0.272$, $\Omega_\Lambda = \Lambda_0/(3H_0^2) = 0.728$, $\Omega_{\rm b} = 0.045$, $h \equiv H_0/(100\ {\rm km\ s^{-1}\ Mpc^{-1}}) = 0.702$, and $\sigma_8 = 0.82$, where the symbols take their usual meaning. We allow 9 levels of additional refinement above the high-resolution region at $\ell = 14$, taking the maximum level in the AMR hierarch to $\ell = 23$ and a formal spatial resoltion of $\Delta x \simeq 12$ pc h^{-1} comoving in our 100 Mpc h^{-1} comoving volume. This corresponds to a dynamical range of $\sim 8 \times 10^6$, higher than that achieved in the Illustris simulation (Vogelsberger et al., 2014)[3]. We allow star formation in the high-resolution region only, where for testing purposes, we set the star formation criteria, defined in sections 2.7.1 and 2.7.2, to:

$$\begin{aligned} n_* &\geq 1\ {\rm cm^{-3}} \\ T_0 &= 2 \times 10^4\ {\rm K}. \end{aligned} \tag{5.25}$$

In the interest of simplicity while developing our methodology, these parameters have not undergone rigorous calibration at this time, however, are calculated using the method set out in section 2.7.2.

Dark matter (sub)-haloes are catalogued using a modified version of the ROCKSTAR phase-space halo finder (Behroozi et al., 2013b), to support multi-resolution particle data-sets. Particles are first divided into 3D Friends-of-Friends (FOF) groups, which are analysed in 6D phase-space to give robust, grid- and shape-independent haloes. Haloes that are not gravitationally bound are removed and properties are computed using the virial overdensity definition of Bryan and Norman (1998).

We perform three simulations with: v_{bc} and our convolved bias parameter (v_{bc} + bias), the v_{bc} field only (v_{bc}), and a case where we initialise the baryon velocity field to be equal to the CDM velocity ($v_{bc} = 0$).

[3]Note, we achieve far higher dynamical range here than in the previous two chapters due to the lack of computational expense associated with radiative transfer.

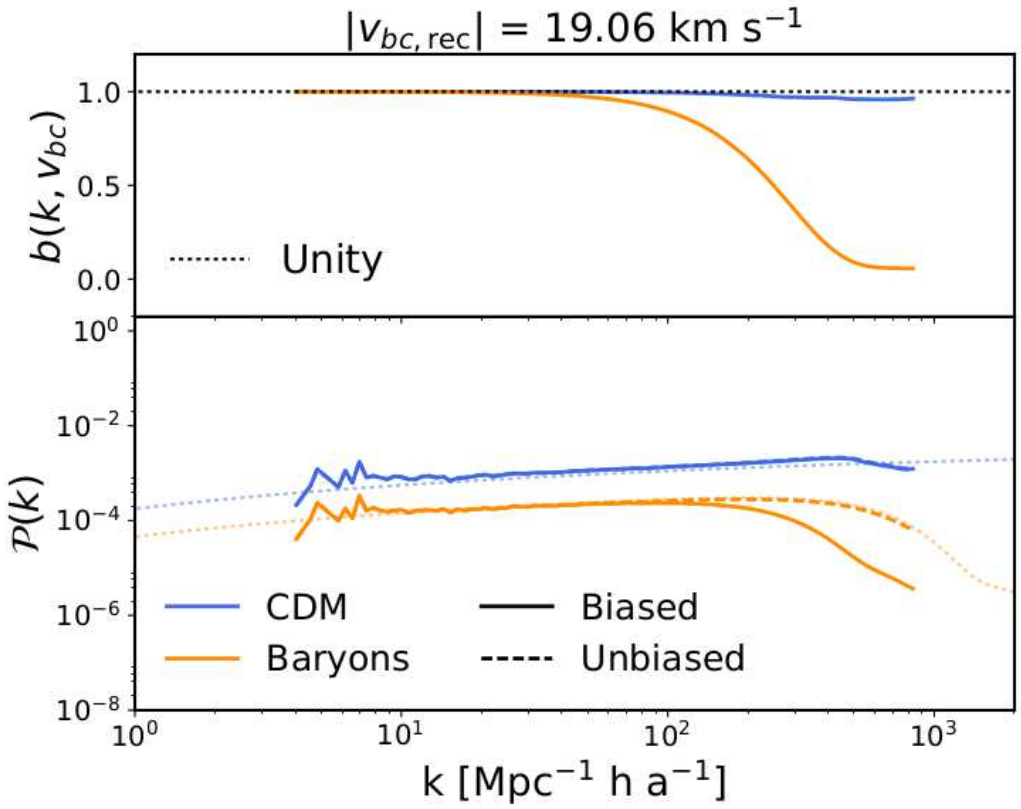

Figure 5.8: *Bottom panel*: Dimensionless power spectra of CDM (blue) and baryons (orange) measured from our high-resolution region with and without convolution of our bias parameter (solid and dashed, respectively), where the dotted lines show the CAMB power spectra as before, where the rms v_{bc} at recombination is $v_{bc,\mathrm{rec}} = 19.06$ km s^{-1} for the whole region. *Top panel*: Globally averaged scale-dependent bias measured by taking the ratio of the biased and unbiased power spectra. For clarity, we draw a line along unity (black dotted). The suppression in the CDM power spectrum is marginal in comparison with baryons, where power is strongly suppressed at small-scales.

5.3 Results

5.3.1 Evolution of the Power Spectrum

To compute the baryon power spectrum from our simulations, we point sample the AMR hierarchy onto regular Cartesian grids, using 512^3 grid cells. As we use Fast-Fourier transforms (FFTs) to compute the power spectrum, we rely on uniform Cartesian grids, and therefore limited in the resolution at which we can sample the density field. This leads to a loss of small-scale power, which cannot be recovered trivially, and would instead require a multi-grid method of power spectrum estimation (or real space correlation function).

As shown in figure 5.1, the geometry of the central high-resolution region is highly ellipsoidal, which leads to complications in point-sampling of the density field. In figure 5.9, we show the gas density field in our central region. The transition from low to high-resolution is clear, as we resolve the intricate web of filaments and first non-linear haloes in the inner-most zoom region only. To compute the baryon matter power spectrum, we draw a cubic region around this inner-most region, with width equal to the maximal radius of the ellipsoid, where we point sample the density field. At the time of writing, this is not possible for dark matter using our CIC smoothing kernel due to significant discreteness effects. A more advanced, multi-grid smoothing kernel is therefore required (however beyond the scope of this initial preliminary work).

We show the estimated baryon power spectra at the initial point of the simulation $z = 200$ (left panel) and $z = 21.21$ (right panel) in figure 5.10 for all three simulations. In the left-hand panel, we clearly see that the modulation of the power spectrum by our scale-dependent bias parameter is captured in RAMSES, while all three simulations converge at (relative) low-k with the linear theory CAMB prediction (black dotted line). In contrast, in the right-hand panel our v_{bc} + bias simulation shows reduced power at low-k relative to our other two simulations, for which there is also a minor discrepancy due to the inclusion of the v_{bc} field without our bias. This is due to non-linear mode coupling, as within the ΛCDM model, hierarchical collapse occurs from low to high-k. Therefore, suppression of high-k modes reduces the growth rate of low-k modes under gravity. The power spectrum is non-linear in all three simulations at $k \lesssim 10$ Mpc^{-1} h a^{-1}.

5.3.2 The Halo Mass Function and Star Formation Rate

In cosmology, the halo mass function measures the abundance of dark matter haloes in the Universe, which is sensitive to the variance of the power spectrum defined in equation 1.24

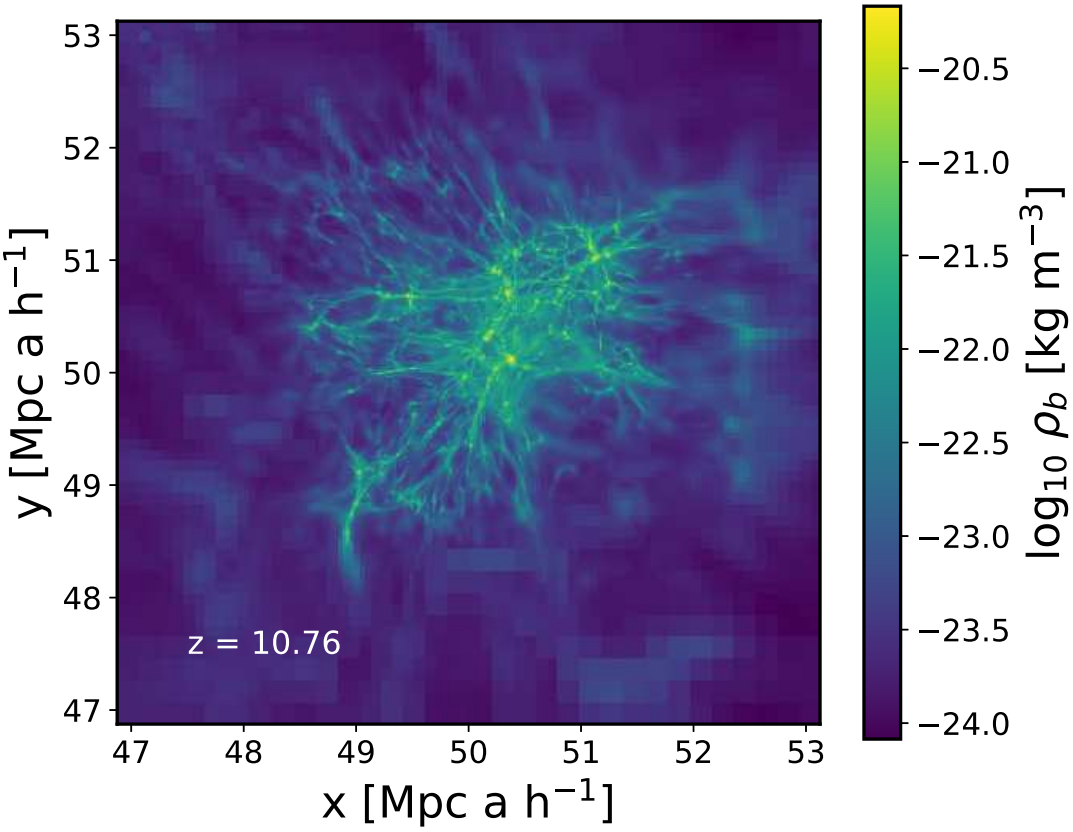

Figure 5.9: Peak gas density along the line of sight through the central high-resolution region in our simulation with biased power spectrum at $z = 10.76$. The transition from low to high resolution is clearly seen, as we reach a formal resolution of $\Delta x \simeq 12$ pc h^{-1} comoving in the high-resolution region.

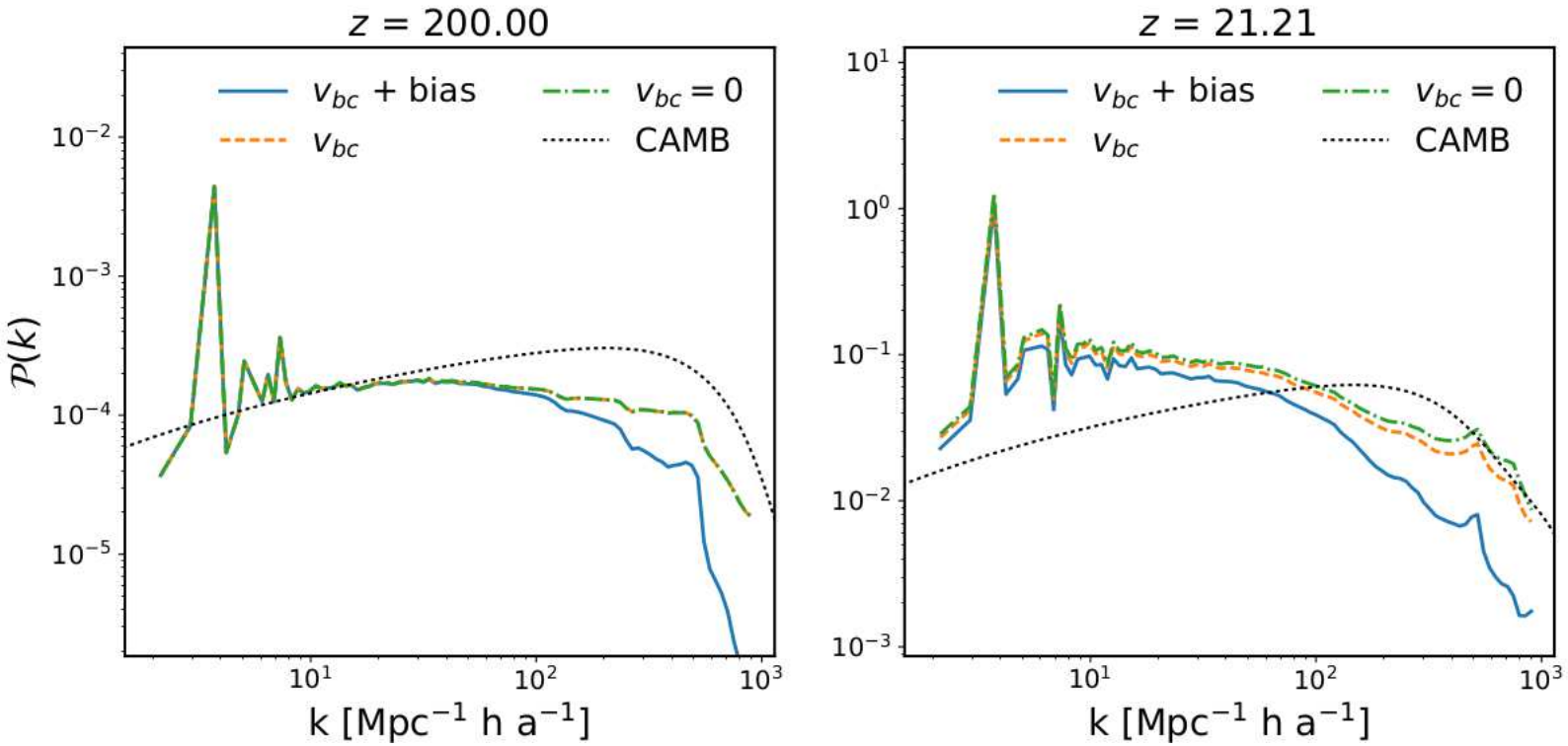

Figure 5.10: Dimensionless power spectra of the baryon density field, measured from our RAMSES simulations, at $z = 200$ (left; initial conditions) and $z = 21.21$ (right), for the high-resolution region (and nearby buffer regions) only. We show three simulations, where we include the drift velocity with and without our convolved bias parameter (v_{bc} + bias and v_{bc}, respectively) and a simulation where baryons are initialised using the CDM velocity field (i.e $v_{bc} = 0$). We show the CAMB power spectra for reference, however direct the reader to the text for an explanation for the discrepancy with our measured power spectra in both panels. At $z = 21.21$, power is suppressed on all scales and interestingly on scales larger than those where our bias is significant due to non-linear mode coupling.

via Press-Schecter theory (Press and Schechter, 1974):

$$\frac{dn}{d\ln M} = \sqrt{\frac{2}{\pi}} \frac{\rho_M}{M} \frac{d\ln\sigma^{-1}}{d\ln M} \nu \, \exp(-\nu^2/2), \tag{5.26}$$

where $\nu \equiv \delta_c/\sigma(M)$ and $dn/d\ln M$ denotes the differential number density of haloes in logarithmic bins of mass. We show the halo mass function from our simulations in figure 5.11 at $z = 15.66$ (left panel) and $z = 10.11$ (right panel). We show the theoretical Sheth-Tormen (ST) prediction (Sheth and Tormen, 2002, grey dashed line) for reference, however note that our region is biased towards high masses by choice.

Our bias parameter leads to a modulation in the mass function at all masses measured at $z = 15.66$. This is due to the suppression in power near the Jeans scale, while our v_bc simulation without bias traces that of our $v_bc = 0$ model. By $z = 10.11$, all three simulations converge, as v_bc decays as $(1 + z)$ and thus becomes negligible by $z \sim 10$. Haloes below the atomic cooling scale, i.e $M \sim 10^8 \ M_\odot \ h^{-1}$, are thought to host the first stars and galaxies, which in turn are predicted to drive the process of cosmic reionization. Modulation of their abundance and growth could have significant implications for the early stages and duration of reionization, and the source population responsible for this process (e.g Barkana and Loeb, 2001; Bromm and Larson, 2004; Wise et al., 2014)

The modulation of the halo mass function comes at a crucial time in the formation of the first stars, which could delay Population III star formation (e.g Maio et al., 2011; Stacy et al., 2011; Greif et al., 2011; O'Leary and McQuinn, 2012). In figure 5.12, we show the star formation rate (SFR) computed using all haloes in our zoom region over the redshift range $22 \lesssim z \lesssim 10$. The SFR is suppressed at early times in our simulation with scale-dependent bias, while the other two models show reasonably good agreement, in part due to the suppression of the halo mass function itself.

At $z \sim 14$, we see that the SFR of our biased simulation peaks, while the remaining two models experience a sudden drop due to SNe feedback. As the SFR has been largely suppressed at early times in our biased model, there are far fewer SNe events recorded which allows the SFR to continue to climb in comparison with our remaining models. Over the range $14 \lesssim z \lesssim 10$, the three SFR histories converge as haloes become self-regulating under SNe feedback and v_bc decays, transitioning to sub-sonic. This would suggest that the impact of the supersonic relative drift velocity on the SFR of haloes could delay the beginning of reionization by an estimated $\Delta z \sim 2$ in redshift (e.g Bittner and Loeb, 2011).

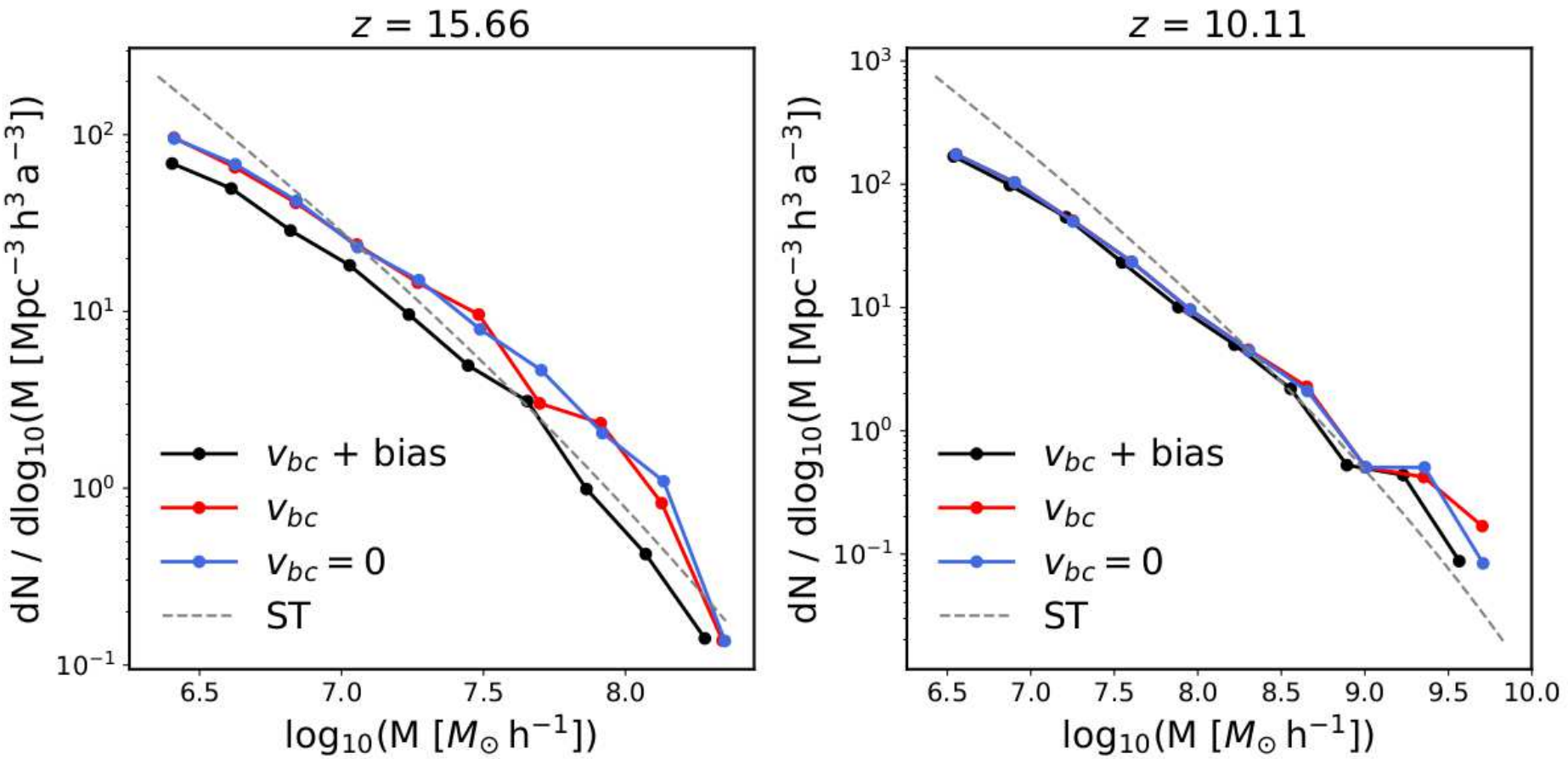

Figure 5.11: Halo mass function at $z = 15.66$ (left panel) and $z = 10.11$ (right panel) in the high-resolution region for all three simulations as indicated. We show the Sheth-Tormen (ST Sheth and Tormen, 2002) for reference (grey dashed line), however note that our zoom region is by construction, chosen to be biased towards high densities and is therefore not representative of the cosmic mean. Convolution of our bias parameter suppresses the halo mass function at early times at all mass scales, crucially when the first stars are beginning to form. By $z \sim 10$, as the drift velocity redshifts away, the mass function of all three simulations begins to converge.

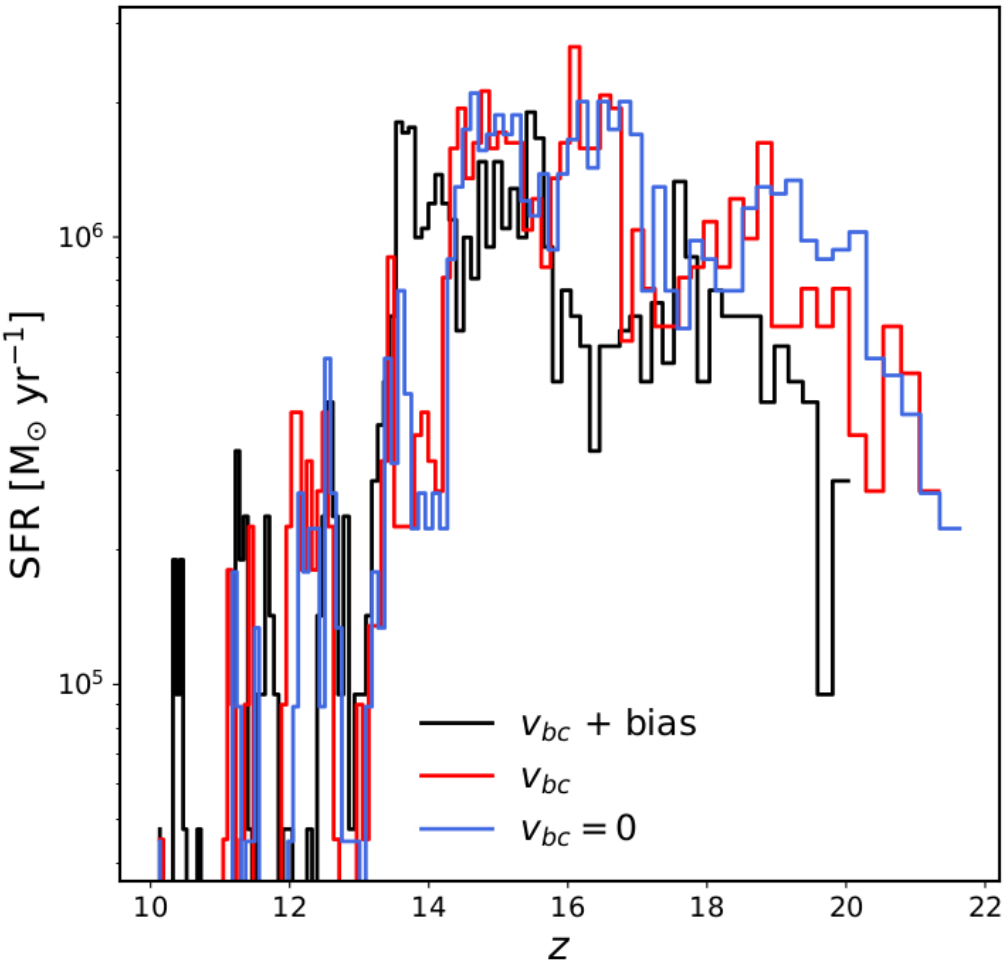

Figure 5.12: Star formation rate of haloes in our zoom region, over the redshift range $22 \lesssim z \lesssim 10$. Modulation of the halo mass function at early times leads to a reduced star formation rate, as baryons are unable to condense into dark matter haloes to form the first stars. Over time, haloes become self-regulating due to SNe feedback, reducing the discrepancy between our three simulations.

5.4 Discussion and Conclusions

We have presented preliminary results for the first large-scale simulations with a relative drift velocity, in volumes an order of magnitude larger on-a-side than the coherence scale of v_{bc}. This effect, which is second-order in the Euler equation of fluid dynamics, modulates small-scale density fluctuations in baryons, and to a lesser but still significant extent dark matter. We modulate the power spectrum of density fluctuations using a scale-dependent bias parameter, which is convolved into the initial linear theory density fields of baryons and CDM. Corrections to the velocity field are also computed, via the Fourier transform of the Continuity equation and applied in our simulations.

We find that convolution of this bias into our initial conditions leads to a modulation of the halo mass function and global star formation rate, in good agreement with previous studies. We stress, that due to the assumption that each sub-region has the mean density of the Universe, our results are likely to over-estimate of the total modulation of the power spectrum. In contrast, our lack of a detailed treatment of how gas pressure modules the growth of modes may lead to an underestimation in the suppression of small-scale modes. Therefore, a detailed study of these effects is required to fully understand our results.

The ability to capture the effects of the relative drift velocity in volumes much larger than the coherence scales, opens up the possibility to fully sample environmental effects on the halo and galaxy population. Our method can be used to re-simulate several sub-volumes from large cosmological simulations with a relative drift velocity for the first time, capturing what could be a crucial feedback mechanism is the previously assumed cold IGM during the dark ages.

The supersonic relative drift velocity between baryons and dark matter, could have significant implications for the formation of the first non-linear objects. The first simulations to fully and self-consistently account for this effect, performed by O'Leary and McQuinn (2012), showed that simply boosting v_{bc} at the initial simulation redshift, without accounting for its impact on the growth of modes, as done previously (e.g Maio et al., 2011; Stacy et al., 2011; Greif et al., 2011; Naoz et al., 2012), misses much of the impact on the linear growth of density fluctuations. In this work, we use the CICSASS code (O'Leary and McQuinn, 2012) to solve the set of equations listed in equation 5.4, for sub-regions below the coherence length of v_{bc}. We defined the scale-dependent bias parameter, which we convolve into our linear-theory density fields to modulate small-scale power.

Our results show, that by including the impact of the drift velocity on the power spectrum, by means of convolution with our bias parameter, the baryon power spectrum

and dark matter halo mass function are both suppressed at early times. The impact of v_{bc} on the halo mass function was first investigated by Naoz et al. (2012), who found that the total number density of haloes is suppressed by $\sim 20\%$ at $z = 15$. We find a higher suppression of $\sim 30\%$ between $16 < z < 14$ in our v_{bc} + bias simulation, which is potentially due to the over modulation of the power spectrum due to our assumption that each sub-region has mean density. By $z = 10.11$, the total number of dark matter haloes is reduced by just $\sim 5\%$, however this could still have significant implications for the duration of reionization.

The global SFR is suppressed due to the modulation of small-scale power and the halo mass function by v_{bc}, which could have important ramification for the first stars. Further modulation of the growth of modes due to gas pressure, stressed to be important by Naoz and Barkana (2005, 2007), could lead to additional suppression in their formation, as found by previous studies (e.g Maio et al., 2011; Stacy et al., 2011; Greif et al., 2011; O'Leary and McQuinn, 2012; Fialkov et al., 2012).

When computing the power spectrum in sub-regions with $v_{bc} > 0$, we assume each patch has mean cosmological density. In practice, this assumption breaks down, as each patch almost never has the mean density. Ahn (2016) recently investigated how the environment changes the way the drift velocity influences the growth of modes, leading to a new type of coupling between large- scale and small-scale modes. In their formalism, variation in the density of the environment, as well as the relative velocities, is accounted for self-consistently. They subsequently show that this leads to a net boost in the overall amplitude of the power spectrum, compared with the prediction by Tseliakhovich and Hirata (2010). This suggests that our bias parameter *over* estimates the suppression of small-scale density modes, which is somewhat negated in our simulations, as we do not capture the full range of large-scale velocity modes in our box, along with the impact of gas pressure. To account for this, we plan to adopt the formalism of Ahn (2016) in the future, to compute the biased and unbiased power spectra of dark matter and baryons.

Although our volume marks a significant step in including the drift velocity in numerical simulations, we fall short of including the BAOs, which source significant v_{bc}. Our method is not technically limited in terms of volume, however the generation of zoom initial conditions, with sufficient resolution to resolve extremely high-k modes, poses computational challenges in terms of available memory. Currently, we use the MUSIC code (Hahn and Abel, 2013) to compute our initial (linear) density and velocity fields. The code crucially supports 2LPT, allowing us to recover small-scale power as shown by the

extremely good agreement with CAMB at the highest-k in our box (figures 5.3 and 5.4). One drawback is that the MUSIC code does not currently support MPI or similar parallel environments which utilise distributed memory clusters. Were such improvements to be implemented, we could generate much larger volumes whilst still resolving high-k modes where our method can be applied.

Acknowledgements

The authors contributed in the following way to this paper. DS undertook the calibration and completion of all simulations; the development of all analysis codes, the production of all figures, and the first draft of the paper. AF helped with the theoretical development of our methodology, and provided expertise in the subject area. II supervised the project and helped with the interpretation of results and feedback on the first draft.

DS would like to thank Joakim Rosdahl for valuable technical feedback, and Alexander Eggemeier for useful discussions and insight.

This work used the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The DiRAC system is funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grant ST/H008519/1, STFC DiRAC Operations grant ST/K003267/1, and Durham University. DiRAC is part of the National E-Infrastructure. We acknowledge PRACE for awarding us computational time under project PRACE4LOFAR grant 2014102339 and "Multi-scale simulations of Cosmic Reionization" grants 2014102281, 2015122822, and 2016153528 to resource Curie based in France at CEA and to resource Marenostrum based in Spain at BSC. This work was supported by the Science and Technology Facilities Council [grant numbers ST/F002858/1 and ST/I000976/1] and the Southeast Physics Network (SEPNet). Some of the analysis was done on the Apollo cluster at The University of Sussex.

We made extensive use of the PYMSES (Guillet et al., 2013) and PYNBODY (Pontzen et al., 2013) python modules when developing our own analysis pipeline.

Chapter 6

Conclusions

In this thesis, we have presented a number of predictions for the properties of galaxies during and prior to the EoR. This was achieved using a suite of cosmological radiation hydrodynamics simulations to study the EoR, and zoom simulations to probe suppression on small-scales, due to the relative drift velocity between dark matter and baryons following recombination. Our simulations represent some of the highest-resolution models produced to date, reaching the tens of parsec scales. In this work, we have focused on the details of feedback and the implications for the first non-linear objects in the Universe. In this chapter, we give a brief summary of our key results and present avenues for future work.

6.1 Summary of Results

In Chapter 3, we focused on the transport of gas from the IGM into haloes during reionization. As the UV background rises, haloes with virial temperatures below that of the photoheated IGM, undergo photoevaporation on time-scales of a few Myr, reducing their baryon fractions to a few percent or less. The main focus of this work was to provide a new intelligent alternative to existing models which constrain this quantity for use in semi-analytical models of galaxy formation.

We show that radiative feedback inhibits the formation of the first stars and galaxies, by preventing the otherwise efficient collapse of gas into haloes. The magnitude of suppression is stronger the earlier reionization begins, as the mass of haloes is, on average, smaller at early times, and therefore more susceptible to feedback. In theory, haloes below the atomic cooling limit are unable to accrete gas once H II regions sweep through their local environment, starving the central galaxy of star forming gas. In practice, we show that while this is largely the case, the baryon fraction of these haloes is sensitive to additional

environmental effects, such as tidal stripping.

As several parameters and physical processes influence the baryon fraction, it is difficult to construct a simple analytical formalism to fit this quantity. Therefore, we developed and trained an artificial neural network, which is better able to capture the full parameter space. In doing so, we demonstrated that our model is independent of redshift, mass, and most importantly, reionization history.

In Chapter 4, we investigated the impact that binary star populations have on the escape fraction of ionizing photons. This parameter is crucial in any model of reionization, however it is both poorly defined and understood. Our work was motivated by Ma et al. (2016), who found that processes such as binary mass transfer, which increase the UV luminosity of older stellar populations, could boost the escape fraction to the levels required by semi-analytical models.

A crucial ingredient missing from prior semi-analytical models, is the negative feedback effect which binary stars exhibit on their host galaxies. Using fully-coupled simulations, we have shown that binaries boost radiative feedback marginally, however significantly enough to negate much of the expected increase in the number of ionizing photons which escape galaxies. The additional heating, due to the hardening of their UV spectra, raises the cosmological Jeans mass (known as the filtering mass), in comparison with our single stellar population model. Although the time-integrated photon escape fraction is raised in our binary star models, we find that the increase is much lower than previously predicted. Our main conclusion therefore, is that including binaries is not a sufficient argument for boosting the UV escape fraction in semi-analytical models.

Finally, in Chapter 5, we presented a new method for including the supersonic relative drift velocity, between baryons and CDM, and its modulation of small-scale density fluctuations in large-scale volumes. Our method relies on the calculation of a scale-dependent bias parameter, in sub-regions where the drift velocity is coherent, such that the Euler equations can be linearised in terms of v_{bc}.

To capture scales where the power spectrum is suppressed, we used cosmological zoom simulations, to resolve a small region at extremely high resolution. Although we do not yet account for the impact of gas pressure on the growth of modes, we presented preliminary results from simulations with and without our bias included. We have shown, in agreement with previous studies, that the modulation of power leads to a suppression in both the mass function of dark matter haloes, and the global star formation rate at $z > 10$. In the absence of a drift velocity, and significant shock heating via structure formation, the IGM

is expected to be kinematically cold during this era, however any supersonic flow would potentially generate entropy.

Any modulation in the formation of the first dark matter haloes and stars could have important implications for the EoR, with some authors predicting a delay by as much as $\Delta z \sim 2$ in its completion. Our method enables radiative transfer simulations of this epoch, accounting for the drift velocity in the process.

In summary, the primary conclusion of this thesis is that both radiative feedback and the supersonic drift velocity play a crucial role in the formation and evolution of the first stars and galaxies, and their effects must be included to produce a complete model of galaxy formation.

6.2 Future Work

In this section, we present some potential future avenues of work based on the work presented thus far.

6.2.1 Halo Baryon Fraction

While our work on the halo baryon fraction marks a significant improvement simple mass-redshift dependent fits, there are avenues for improvement. Firstly, our volumes are relatively small due to computational requirements. Therefore, we do not capture to high-mass halo population and much of the statistical variation in the history and progression of reionization. It would be both useful and interesting to train our neural network on much larger volumes to test the robustness of our results. Furthermore, it would be interesting to run our simulations to $z = 0$ to investigate whether our results converge with those of prior authors at the present day.

An additional avenue for future work lies in training more sophisticated neural networks with additional hidden layers (deep learning). As our network is quite simplistic in nature, using a single hidden layer (shallow learning), there is a high chance that more sophisticated networks could improve the accuracy of our results. The same outcome could also be achieved by adding additional parameters (degrees of freedom) to our model, however this has the drawback that additional quantities be constrained in order to use the model predictions.

6.2.2 Escape Fraction

As above, our simulations suffer from poor statistics due to their limited volume. This is especially problematic for understanding how the escape fraction varies at higher halo masses, which would be especially interesting for Milky Way sized haloes. This would also allow us to better compare our UV luminosity functions with observations.

Low mass haloes below the atomic cooling limit would also be interesting to study, especially as these mini-haloes are expected to drive cosmic reionization. A detailed treatment of molecular chemistry would be required to include these haloes, as their primary cooling mechanism is the vibrational and rotational lines of molecular hydrogen.

Kimm and Cen (2014) have shown that the choice of SNe feedback model and inclusion of runaway OB stars could significantly impact the escape fraction. A detailed study of various SNe feedback models would therefore be interesting in the context of binary versus single stellar population modelling.

6.2.3 Relative Drift Velocity

In Chapter 5, we have already discussed potential avenues for future work, such as including the impact of gas pressure on the growth of modes and accounting for the variation in environment with respect to the mean density of each subregion.

It would be of great interest to follow up our findings for the halo baryon fraction with the relative drift velocity included. While our simulation volumes presented in Chapter 3 are relatively small, below the coherence length where a coherent flow can be placed across the entire box, our new method would also allow us to improve upon our statistics. Previous studies have already investigated the impact of a supersonic relative drift velocity on the halo baryon fraction (e.g Naoz et al., 2013), therefore this would present a natural follow up of both Chapters 3 and 5.

An additional highly interesting avenue would be in investigating what impact, if any, the coherence supersonic flow could have weak lensing statistics. A major contaminant of weak lensing studies is the intrinsic alignment of galaxies along the axis of the local tidal field. As the drift velocity is coherent over several Mpc comoving, this effect could also introduce intrinsic alignment effects. Intrinsic alignments contaminate the real space two-point correlation function of galaxies, $\xi(r)$, which is defined as the sum of all potential alignment scenarios, corresponding to shear, intrinsic-intrinsic (II), galaxy-intrinsic (GI) and intrinsic-galaxy (IG):

$$\xi(r) = \langle \epsilon_i^{obs} \epsilon_j^{obs} \rangle = \langle \gamma_i \gamma_j \rangle + \langle \epsilon_i^{int} \epsilon_j^{int} \rangle + \langle \gamma_i \epsilon_j^{int} \rangle + \langle \epsilon_i^{int} \gamma_j \rangle, \tag{6.1}$$

where γ_i denotes the complex shear field.

Our method could be used to compute contributions from the drift velocity to the $\langle \epsilon_i^{int} \epsilon_j^{int} \rangle$ term, by re-simulating several zoom regions in our box to capture the full shear statistics.

Chapter 7

Publications

Below I have included references to manuscripts of which I have co-authored during my PhD, with a brief outline of their conclusions and my contribution.

Ocvirk, P., Gillet, N., Shapiro, P. R., Aubert, D., Iliev, I. T., Teyssier, R., Yepes, G., Choi, J.-H., Sullivan, D., Knebe, A., Gottlo ber, S., DAloisio, A., Park, H., Hoffman, Y., and Stranex, T. (2016). Cosmic Dawn (CoDa): the First Radiation-Hydrodynamics Simulation of Reionization and Galaxy Formation in the Local Universe. MNRAS, 463:14621485.

In this paper, we performed the largest, fully-coupled, cosmological radiation hydrodynamical simulations of galaxy formation during the EoR. It was performed on Titan at Oak Ridge National Laboratory using RAMSES-CUDATON deployed on 8192 nodes, using 1 core and 1 GPU per node. This is the first time a GPU-accelerated, fully coupled RHD code has been used on such a scale. It was carried out in collaboration with the CLUES[1] project, who provided the initial conditions which are constrained to reproduce the local group of galaxies.

The primary aim of this project was to investigate the morphology of reionization in our local group, exploring whether the Milky Way was ionized from the inside-out or outside-in, potentially by the local Virgo supercluster. This paper served as the initial methodology paper, in which we presented some basic results from the simulation, such as reionization history. We chose to focus on self-feedback primarily, drawing the conclusion that radiative feedback is potentially far more efficient at suppressing star formation than previously thought.

My contribution to this work involved weekly Skype calls to discuss the first res-

[1]Constrained Local UniversE Simulations, http://www.clues-project.org/index.html.

ults and direction of the paper, re-simulation of a small calibration volume on which to develop post-processing tools and check our results, and finally understanding and comparing our halo mass functions with those produced using pure N-body simulations by the CLUES project. This turned out to be crucial, as the initial halo catalogue produced from the CoDa simulation suffered from over-linking at high-redshifts; a common issue with Friends-of-Friends halo finders. The halo catalogues were ultimately reproduced using a lower linking length at high-redshift, which corrected our results.

Bisbas, T. G., Haworth, T. J., Williams, R. J. R., Mackey, J., Tremblin, P., Raga, A. C., Arthur, S. J., Baczynski, C., Dale, J. E., Frostholm, T., Geen, S., Haugbølle, T., Hubber, D., Iliev, I. T., Kuiper, R., Rosdahl, J., Sullivan, D., Walch, S., and Wünsch, R. (2015). STARBENCH: the D-type expansion of an H II region. MNRAS, 453:13241343

STARBENCH is a project focused on benchmarking and validating different star-formation and stellar feedback codes. In this first STARBENCH paper we perform a comparison study of the D-type expansion of an H II region. I attended the first STARBENCH workshop at the Argelander-Institut für Astronomie, in Bonn, aimed at the collation and discussion of initial results. In collaboration with Joakim Rosdahl and Sam Geen, I took part in the testing of both RAMSES-RT and RAMSES-CUDATON, of which only the former was included in the final work.

Since the codes do not agree with any of the analytic expressions presented, we developed an analytic parametrization that describes the early and late phase expansion of an H II region to within 2% at all times. This parametrization is recommend to be used for future code validations, exercises and analytical studies of massive star-forming regions.

Bibliography

Abel, T., Anninos, P., Zhang, Y., and Norman, M. L. (1997). Modeling primordial gas in numerical cosmology. New A, 2:181–207. 25

Agertz, O., Kravtsov, A. V., Leitner, S. N., and Gnedin, N. Y. (2013). Toward a Complete Accounting of Energy and Momentum from Stellar Feedback in Galaxy Formation Simulations. ApJ, 770:25. xi, 25, 26

Agertz, O., Moore, B., Stadel, J., Potter, D., Miniati, F., Read, J., Mayer, L., Gawryszczak, A., Kravtsov, A., Nordlund, Å., Pearce, F., Quilis, V., Rudd, D., Springel, V., Stone, J., Tasker, E., Teyssier, R., Wadsley, J., and Walder, R. (2007). Fundamental differences between SPH and grid methods. MNRAS, 380:963–978. 31

Aghanim, N., Majumdar, S., and Silk, J. (2008). Secondary anisotropies of the CMB. *Reports on Progress in Physics*, 71(6):066902. 22

Ahn, K. (2016). How the Density Environment Changes the Influence of the Dark Matter-Baryon Streaming Velocity on Cosmological Structure Formation. ApJ, 830:68. 153

Ahn, K., Shapiro, P. R., Iliev, I. T., Mellema, G., and Pen, U.-L. (2009). The Inhomogeneous Background Of H_2-Dissociating Radiation During Cosmic Reionization. ApJ, 695:1430–1445. 27

Alvarez, M. A. and Abel, T. (2007). Quasar HII regions during cosmic reionization. MNRAS, 380:L30–L34. 88

Arribas, S., Colina, L., Bellocchi, E., Maiolino, R., and Villar-Martín, M. (2014). Ionized gas outflows and global kinematics of low-z luminous star-forming galaxies. A&A, 568:A14. 44

Aubert, D., Deparis, N., and Ocvirk, P. (2015). EMMA: an adaptive mesh refinement cosmological simulation code with radiative transfer. MNRAS, 454:1012–1037. 81, 83

Aubert, D. and Teyssier, R. (2008). A radiative transfer scheme for cosmological reionization based on a local Eddington tensor. MNRAS, 387:295–307. 31, 32, 34, 37, 39, 57, 89

Aubert, D. and Teyssier, R. (2010). Reionization Simulations Powered by Graphics Processing Units. I. On the Structure of the Ultraviolet Radiation Field. ApJ, 724:244–266. 57, 92

Babul, A. and Rees, M. J. (1992). On dwarf elliptical galaxies and the faint blue counts. MNRAS, 255:346–350. 20, 55

Baertschiger, T. and Sylos Labini, F. (2002). On the problem of initial conditions in cosmological N-body simulations. *EPL (Europhysics Letters)*, 57:322–328. 34

Barkana, R. and Loeb, A. (1999). The Photoevaporation of Dwarf Galaxies during Reionization. ApJ, 523:54–65. 27

Barkana, R. and Loeb, A. (2000). Identifying the Reionization Redshift from the Cosmic Star Formation Rate. ApJ, 539:20–25. 86

Barkana, R. and Loeb, A. (2001). In the beginning: the first sources of light and the reionization of the universe. Phys. Rep., 349:125–238. 19, 97, 149

Becker, R. H., Fan, X., White, R. L., Strauss, M. A., Narayanan, V. K., Lupton, R. H., Gunn, J. E., Annis, J., Bahcall, N. A., Brinkmann, J., Connolly, A. J., Csabai, I., Czarapata, P. C., Doi, M., Heckman, T. M., Hennessy, G. S., Ivezić, Ž., Knapp, G. R., Lamb, D. Q., McKay, T. A., Munn, J. A., Nash, T., Nichol, R., Pier, J. R., Richards, G. T., Schneider, D. P., Stoughton, C., Szalay, A. S., Thakar, A. R., and York, D. G. (2001). Evidence for Reionization at z˜6: Detection of a Gunn-Peterson Trough in a z=6.28 Quasar. AJ, 122:2850–2857. 22, 98

Behroozi, P. S., Wechsler, R. H., and Conroy, C. (2013a). The Average Star Formation Histories of Galaxies in Dark Matter Halos from z = 0-8. ApJ, 770:57. xiii, xvi, 66, 68, 70, 116

Behroozi, P. S., Wechsler, R. H., and Wu, H.-Y. (2013b). The ROCKSTAR Phase-space Temporal Halo Finder and the Velocity Offsets of Cluster Cores. ApJ, 762:109. 52, 58, 90, 103, 144

Behroozi, P. S., Wechsler, R. H., Wu, H.-Y., Busha, M. T., Klypin, A. A., and Primack, J. R. (2013c). Gravitationally Consistent Halo Catalogs and Merger Trees for Precision Cosmology. ApJ, 763:18. 58, 103, 113

Belokurov, V., Zucker, D. B., Evans, N. W., Gilmore, G., Vidrih, S., Bramich, D. M., Newberg, H. J., Wyse, R. F. G., Irwin, M. J., Fellhauer, M., Hewett, P. C., Walton, N. A., Wilkinson, M. I., Cole, N., Yanny, B., Rockosi, C. M., Beers, T. C., Bell, E. F., Brinkmann, J., Ivezić, Ž., and Lupton, R. (2006). The Field of Streams: Sagittarius and Its Siblings. ApJ, 642:L137–L140. 68

Benson, A. J. (2010). Galaxy formation theory. Phys. Rep., 495:33–86. 25

Benson, A. J., Frenk, C. S., Lacey, C. G., Baugh, C. M., and Cole, S. (2002a). The effects of photoionization on galaxy formation - II. Satellite galaxies in the Local Group. MNRAS, 333:177–190. 55

Benson, A. J., Lacey, C. G., Baugh, C. M., Cole, S., and Frenk, C. S. (2002b). The effects of photoionization on galaxy formation - I. Model and results at z=0. MNRAS, 333:156–176. 55

Bergvall, N., Leitet, E., Zackrisson, E., and Marquart, T. (2013). Lyman continuum leaking galaxies. Search strategies and local candidates. A&A, 554:A38. 98

Bertelli, G., Bressan, A., Chiosi, C., Fagotto, F., and Nasi, E. (1994). Theoretical isochrones from models with new radiative opacities. A&AS, 106. 28

Bertschinger, E. (2001). Multiscale Gaussian Random Fields and Their Application to Cosmological Simulations. ApJS, 137:1–20. 34, 133

Bittner, J. M. and Loeb, A. (2011). The imprint of the relative velocity between baryons and dark matter on the 21-cm signal from reionization. *ArXiv e-prints*. 149

Bond, J. R. and Efstathiou, G. (1984). Cosmic background radiation anisotropies in universes dominated by nonbaryonic dark matter. ApJ, 285:L45–L48. 22

Boutsia, K., Grazian, A., Giallongo, E., Fontana, A., Pentericci, L., Castellano, M., Zamorani, G., Mignoli, M., Vanzella, E., Fiore, F., Lilly, S. J., Gallozzi, S., Testa, V., Paris, D., and Santini, P. (2011). A Low Escape Fraction of Ionizing Photons of L > L* Lyman Break Galaxies at z = 3.3. ApJ, 736:41. 98

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., Trenti, M., Labbé, I., Bradley, L., Carollo, M., van Dokkum, P. G., Gonzalez, V., Holwerda, B., Franx, M., Spitler, L., Smit, R., and Magee, D. (2015). UV Luminosity Functions at Redshifts z $\sim$ 4 to z $\sim$ 10: 10,000 Galaxies from HST Legacy Fields. ApJ, 803:34. xvi, 115, 117, 122

Bromm, V. and Larson, R. B. (2004). The First Stars. ARA&A, 42:79–118. 19, 27, 149

Bruzual, G. and Charlot, S. (2003). Stellar population synthesis at the resolution of 2003. MNRAS, 344:1000–1028. xii, 29, 44, 46, 58, 89

Bruzual A., G. and Charlot, S. (1993). Spectral evolution of stellar populations using isochrone synthesis. ApJ, 405:538–553. 29

Bryan, G. L. and Norman, M. L. (1998). Statistical Properties of X-Ray Clusters: Analytic and Numerical Comparisons. ApJ, 495:80–99. 16, 58, 103, 144

Bryan, G. L., Norman, M. L., O'Shea, B. W., Abel, T., Wise, J. H., Turk, M. J., Reynolds, D. R., Collins, D. C., Wang, P., Skillman, S. W., Smith, B., Harkness, R. P., Bordner, J., Kim, J.-h., Kuhlen, M., Xu, H., Goldbaum, N., Hummels, C., Kritsuk, A. G., Tasker, E., Skory, S., Simpson, C. M., Hahn, O., Oishi, J. S., So, G. C., Zhao, F., Cen, R., Li, Y., and Enzo Collaboration (2014). ENZO: An Adaptive Mesh Refinement Code for Astrophysics. ApJS, 211:19. 40

Calverley, A. P., Becker, G. D., Haehnelt, M. G., and Bolton, J. S. (2011). Measurements of the ultraviolet background at 4.6 $<$ z $<$ 6.4 using the quasar proximity effect. MNRAS, 412:2543–2562. 61, 63, 104

Cantiello, M., Yoon, S.-C., Langer, N., and Livio, M. (2007). Binary star progenitors of long gamma-ray bursts. A&A, 465:L29–L33. 101

Chabrier, G. (2003). Galactic Stellar and Substellar Initial Mass Function. PASP, 115:763–795. 28, 44, 58

Chisholm, J., Tremonti, C. A., Leitherer, C., Chen, Y., Wofford, A., and Lundgren, B. (2015). Scaling Relations Between Warm Galactic Outflows and Their Host Galaxies. ApJ, 811:149. 44

Chornock, R., Berger, E., Fox, D. B., Lunnan, R., Drout, M. R., Fong, W.-f., Laskar, T., and Roth, K. C. (2013). GRB 130606A as a Probe of the Intergalactic Medium and the Interstellar Medium in a Star-forming Galaxy in the First Gyr after the Big Bang. ApJ, 774:26. 63, 104

Choudhury, T. R. and Ferrara, A. (2006). Physics of Cosmic Reionization. *ArXiv Astrophysics e-prints*. 19

Choudhury, T. R., Haehnelt, M. G., and Regan, J. (2009). Inside-out or outside-in: the topology of reionization in the photon-starved regime suggested by Lyα forest data. MNRAS, 394:960–977. 59, 88

Ciardi, B. and Ferrara, A. (2005). The First Cosmic Structures and Their Effects. Space Sci. Rev., 116:625–705. 19

Ciardi, B., Ferrara, A., Governato, F., and Jenkins, A. (2000). Inhomogeneous reionization of the intergalactic medium regulated by radiative and stellar feedbacks. MNRAS, 314:611–629. 19

Ciardi, B., Ferrara, A., Marri, S., and Raimondo, G. (2001). Cosmological reionization around the first stars: Monte Carlo radiative transfer. MNRAS, 324:381–388. 32

Ciardi, B., Ferrara, A., and White, S. D. M. (2003a). Early reionization by the first galaxies. MNRAS, 344:L7–L11. 59, 88

Ciardi, B., Stoehr, F., and White, S. D. M. (2003b). Simulating intergalactic medium reionization. MNRAS, 343:1101–1109. 92

Clark, P. C., Glover, S. C. O., Klessen, R. S., and Bromm, V. (2011). Gravitational Fragmentation in Turbulent Primordial Gas and the Initial Mass Function of Population III Stars. ApJ, 727:110. 74

Cordier, D., Pietrinferni, A., Cassisi, S., and Salaris, M. (2007). A Large Stellar Evolution Database for Population Synthesis Studies. III. Inclusion of the Full Asymptotic Giant Branch Phase and Web Tools for Stellar Population Analyses. AJ, 133:468–478. 29

Couchman, H. M. P. and Rees, M. J. (1986). Pregalactic evolution in cosmologies with cold dark matter. MNRAS, 221:53–62. 55

Creasey, P., Theuns, T., Bower, R. G., and Lacey, C. G. (2011). Numerical overcooling in shocks. MNRAS, 415:3706–3720. 109

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems (MCSS)*, 2(4):303–314. 75

Davis, A. J., D'Aloisio, A., and Natarajan, P. (2011). Virialization of high-redshift dark matter haloes. MNRAS, 416:242–247. 74

Deharveng, J.-M., Buat, V., Le Brun, V., Milliard, B., Kunth, D., Shull, J. M., and Gry, C. (2001). Constraints on the Lyman continuum radiation from galaxies: First results with FUSE on Mrk 54. A&A, 375:805–813. 98, 113

Devriendt, J. E. G., Guiderdoni, B., and Sadat, R. (1999). Galaxy modelling. I. Spectral energy distributions from far-UV to sub-mm wavelengths. A&A, 350:381–398. 29

Dixon, K. L., Iliev, I. T., Mellema, G., Ahn, K., and Shapiro, P. R. (2016). The large-scale observational signatures of low-mass galaxies during reionization. MNRAS, 456:3011–3029. 59, 86

Doroshkevich, A. G., Zel'dovich, Y. B., and Novikov, I. D. (1967). The Origin of Galaxies in an Expanding Universe. Soviet Ast., 11:233. 55

Dubois, Y. and Teyssier, R. (2008). On the onset of galactic winds in quiescent star forming galaxies. A&A, 477:79–94. 43, 44, 55, 110

Efstathiou, G. (1992). Suppressing the formation of dwarf galaxies via photoionization. MNRAS, 256:43P–47P. 20, 55

Eldridge, J. J., Izzard, R. G., and Tout, C. A. (2008). The effect of massive binaries on stellar populations and supernova progenitors. MNRAS, 384:1109–1118. 89, 99, 100, 103

Eldridge, J. J. and Stanway, E. R. (2012). The effect of stellar evolution uncertainties on the rest-frame ultraviolet stellar lines of C IV and He II in high-redshift Lyman-break galaxies. MNRAS, 419:479–489. 101

Engelbracht, C. W., Rieke, G. H., Gordon, K. D., Smith, J.-D. T., Werner, M. W., Moustakas, J., Willmer, C. N. A., and Vanzi, L. (2008). Metallicity Effects on Dust Properties in Starbursting Galaxies. ApJ, 678:804–827. 108

Fan, X., Narayanan, V. K., Lupton, R. H., Strauss, M. A., Knapp, G. R., Becker, R. H., White, R. L., Pentericci, L., Leggett, S. K., Haiman, Z., Gunn, J. E., Ivezić, Ž., Schneider, D. P., Anderson, S. F., Brinkmann, J., Bahcall, N. A., Connolly, A. J., Csabai, I., Doi, M., Fukugita, M., Geballe, T., Grebel, E. K., Harbeck, D., Hennessy, G., Lamb, D. Q., Miknaitis, G., Munn, J. A., Nichol, R., Okamura, S., Pier, J. R., Prada, F., Richards, G. T., Szalay, A., and York, D. G. (2001). A Survey of z>5.8 Quasars in the Sloan Digital Sky Survey. I. Discovery of Three New Quasars and the Spatial Density of Luminous Quasars at z˜6. AJ, 122:2833–2849. 22, 98

Fan, X., Strauss, M. A., Becker, R. H., White, R. L., Gunn, J. E., Knapp, G. R., Richards, G. T., Schneider, D. P., Brinkmann, J., and Fukugita, M. (2006). Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with z˜6 Quasars. II. A Sample of 19 Quasars. AJ, 132:117–136. 21, 22, 63, 104

Fang, T. and Cen, R. (2004). The Transition from Population III to Population II Stars. ApJ, 616:L87–L90. 28

Faucher-Giguère, C.-A., Kereš, D., Dijkstra, M., Hernquist, L., and Zaldarriaga, M. (2010). Lyα Cooling Emission from Galaxy Formation. ApJ, 725:633–657. 103, 105

Faucher-Giguère, C.-A., Lidz, A., Hernquist, L., and Zaldarriaga, M. (2008a). A Flat Photoionization Rate at 2 <= z <= 4.2: Evidence for a Stellar-Dominated UV Background and against a Decline of Cosmic Star Formation beyond z ˜ 3. ApJ, 682:L9. 98

Faucher-Giguère, C.-A., Lidz, A., Hernquist, L., and Zaldarriaga, M. (2008b). Evolution of the Intergalactic Opacity: Implications for the Ionizing Background, Cosmic Star Formation, and Quasar Activity. ApJ, 688:85–107. 97

Faucher-Giguère, C.-A., Lidz, A., Zaldarriaga, M., and Hernquist, L. (2009). A New Calculation of the Ionizing Background Spectrum and the Effects of He II Reionization. ApJ, 703:1416–1443. xii, 45, 46

Feng, Y., Croft, R. A. C., Di Matteo, T., and Khandai, N. (2013). Growth and anisotropy of ionization fronts near high-redshift quasars in the MassiveBlack simulation. MNRAS, 429:1554–1563. 33

Fialkov, A. (2014). Supersonic relative velocity between dark matter and baryons: A review. *International Journal of Modern Physics D*, 23:1430017. 17, 128

Fialkov, A., Barkana, R., Tseliakhovich, D., and Hirata, C. M. (2012). Impact of the relative motion between the dark matter and baryons on the first stars: semi-analytical modelling. MNRAS, 424:1335–1345. 127, 153

Finkelstein, S. L., Papovich, C., Ryan, R. E., Pawlik, A. H., Dickinson, M., Ferguson, H. C., Finlator, K., Koekemoer, A. M., Giavalisco, M., Cooray, A., Dunlop, J. S., Faber, S. M., Grogin, N. A., Kocevski, D. D., and Newman, J. A. (2012). CANDELS: The Contribution of the Observed Galaxy Population to Cosmic Reionization. ApJ, 758:93. 99, 123

Finlator, K., Özel, F., Davé, R., and Oppenheimer, B. D. (2009). The late reionization of filaments. MNRAS, 400:1049–1061. 92

Fioc, M. and Rocca-Volmerange, B. (1999). A statistical study of nearby galaxies. I. NIR growth curves and optical-to-NIR colors as a function of type, luminosity and inclination. A&A, 351:869–882. 29

Fisher, D. B., Bolatto, A. D., Herrera-Camus, R., Draine, B. T., Donaldson, J., Walter, F., Sandstrom, K. M., Leroy, A. K., Cannon, J., and Gordon, K. (2014). The rarity of dust in metal-poor galaxies. Nature, 505:186–189. 108

Fontanot, F., Cristiani, S., and Vanzella, E. (2012). On the relative contribution of high-redshift galaxies and active galactic nuclei to reionization. MNRAS, 425:1413–1420. 98

Furlanetto, S. R. (2006). The global 21-centimeter background from high redshifts. MNRAS, 371:867–878. 128

Furlanetto, S. R. (2016). The 21-cm Line as a Probe of Reionization. *Understanding the Epoch of Cosmic Reionization: Challenges and Progress*, 423:247. 23

Furlanetto, S. R. and Loeb, A. (2004). Large-Scale Structure Shocks at Low and High Redshifts. ApJ, 611:642–654. 17

Furlanetto, S. R., Oh, S. P., and Briggs, F. H. (2006). Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe. Phys. Rep., 433:181–301. 19, 23, 128

Galametz, M., Madden, S. C., Galliano, F., Hony, S., Bendo, G. J., and Sauvage, M. (2011). Probing the dust properties of galaxies up to submillimetre wavelengths. II. Dust-to-gas mass ratio trends with metallicity and the submm excess in dwarf galaxies. A&A, 532:A56. 108

Geen, S., Hennebelle, P., Tremblin, P., and Rosdahl, J. (2015). Photoionization feedback in a self-gravitating, magnetized, turbulent cloud. *Monthly Notices of the Royal Astronomical Society*, 454(4):4484–4502. 99

Geil, P. M. and Wyithe, J. S. B. (2008). The impact of a percolating IGM on redshifted 21-cm observations of quasar HII regions. MNRAS, 386:1683–1694. 59, 88

Girardi, L., Bressan, A., Bertelli, G., and Chiosi, C. (2000). Evolutionary tracks and isochrones for low- and intermediate-mass stars: From 0.15 to 7 M_{sun}, and from Z=0.0004 to 0.03. A&AS, 141:371–383. 28, 29

Gnedin, N. Y. (2000a). Cosmological Reionization by Stellar Sources. ApJ, 535:530–554. 19, 33, 55, 60

Gnedin, N. Y. (2000b). Effect of Reionization on Structure Formation in the Universe. ApJ, 542:535–541. 13, 20, 45, 56, 61, 81, 84

Gnedin, N. Y. and Abel, T. (2001). Multi-dimensional cosmological radiative transfer with a Variable Eddington Tensor formalism. New A, 6:437–455. 31, 39, 59, 89, 103

Gnedin, N. Y. and Hui, L. (1998). Probing the Universe with the Lyalpha forest - I. Hydrodynamics of the low-density intergalactic medium. MNRAS, 296:44–55. 12

Gnedin, N. Y., Kravtsov, A. V., and Chen, H.-W. (2008). Escape of Ionizing Radiation from High-Redshift Galaxies. ApJ, 672:765–775. 98, 123

Gnedin, N. Y. and Ostriker, J. P. (1997). Reionization of the Universe and the Early Production of Metals. ApJ, 486:581–598. 19, 56

Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., and Bartelmann, M. (2005). HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. ApJ, 622:759–771. 105

Greif, T. H., White, S. D. M., Klessen, R. S., and Springel, V. (2011). The Delay of Population III Star Formation by Supersonic Streaming Velocities. ApJ, 736:147. 18, 74, 127, 149, 152, 153

Grissom, R. L., Ballantyne, D. R., and Wise, J. H. (2014). On the contribution of active galactic nuclei to reionization. A&A, 561:A90. 98

Guillet, T., Chapon, D., and Labadens, M. (2013). PyMSES: Python modules for RAMSES. Astrophysics Source Code Library. 51, 91, 125, 154

Gunn, J. E. and Peterson, B. A. (1965). On the Density of Neutral Hydrogen in Intergalactic Space. ApJ, 142:1633–1641. 21

Haardt, F. and Madau, P. (1996). Radiative Transfer in a Clumpy Universe. II. The Ultraviolet Extragalactic Background. ApJ, 461:20. 85, 88, 89

Haardt, F. and Madau, P. (2001). Modelling the UV/X-ray cosmic background with CUBA. In Neumann, D. M. and Tran, J. T. V., editors, *Clusters of Galaxies and the High Redshift Universe Observed in X-rays*. 21, 85, 87, 89

Haardt, F. and Salvaterra, R. (2015). High-redshift active galactic nuclei and H I reionisation: limits from the unresolved X-ray background. A&A, 575:L16. 98

Hahn, O. and Abel, T. (2013). MUSIC: MUlti-Scale Initial Conditions. Astrophysics Source Code Library. 34, 58, 101, 133, 153

Haiman, Z., Thoul, A. A., and Loeb, A. (1996). Cosmological Formation of Low-Mass Objects. ApJ, 464:523. 27

Hanany, S., Ade, P., Balbi, A., Bock, J., Borrill, J., Boscaleri, A., de Bernardis, P., Ferreira, P. G., Hristov, V. V., Jaffe, A. H., Lange, A. E., Lee, A. T., Mauskopf, P. D., Netterfield, C. B., Oh, S., Pascale, E., Rabii, B., Richards, P. L., Smoot, G. F., Stompor, R., Winant, C. D., and Wu, J. H. P. (2000). MAXIMA-1: A Measurement of the Cosmic Microwave Background Anisotropy on Angular Scales of 10'-5°. ApJ, 545:L5–L9. 5

Heath, D. J. (1977). The growth of density perturbations in zero pressure Friedmann-Lemaitre universes. MNRAS, 179:351–358. 10

Heckman, T. M., Borthakur, S., Overzier, R., Kauffmann, G., Basu-Zych, A., Leitherer, C., Sembach, K., Martin, D. C., Rich, R. M., Schiminovich, D., and Seibert, M. (2011). Extreme Feedback and the Epoch of Reionization: Clues in the Local Universe. ApJ, 730:5. 98, 113

Henriques, B. M. B., White, S. D. M., Thomas, P. A., Angulo, R., Guo, Q., Lemson, G., Springel, V., and Overzier, R. (2015). Galaxy formation in the Planck cosmology - I. Matching the observed evolution of star formation rates, colours and stellar masses. MNRAS, 451:2663–2680. 74, 81

Hinshaw, G., Spergel, D. N., Verde, L., Hill, R. S., Meyer, S. S., Barnes, C., Bennett, C. L., Halpern, M., Jarosik, N., Kogut, A., Komatsu, E., Limon, M., Page, L., Tucker, G. S., Weiland, J. L., Wollack, E., and Wright, E. L. (2003). First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: The Angular Power Spectrum. ApJS, 148:135–159. 5

Hockney, R. W. and Eastwood, J. W. (1988). *Computer simulation using particles*. 131

Hoeft, M., Yepes, G., Gottlöber, S., and Springel, V. (2006). Dwarf galaxies in voids: suppressing star formation with photoheating. MNRAS, 371:401–414. 20, 45, 55, 56, 61, 79, 83, 85, 88, 89

Hopkins, P. F. (2015). A new class of accurate, mesh-free hydrodynamic simulation methods. MNRAS, 450:53–110. 119

Hopkins, P. F., Kereš, D., Oñorbe, J., Faucher-Giguère, C.-A., Quataert, E., Murray, N., and Bullock, J. S. (2014). Galaxies on FIRE (Feedback In Realistic Environments): stellar feedback explains cosmologically inefficient star formation. MNRAS, 445:581–603. 119

Hopkins, P. F., Narayanan, D., and Murray, N. (2013). The meaning and consequences of star formation criteria in galaxy models with resolved stellar feedback. MNRAS, 432:2647–2653. xii, 47, 48, 49, 50

Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366. 75

Hu, W. and Sugiyama, N. (1995). Anisotropies in the cosmic microwave background: an analytic approach. ApJ, 444:489–506. 17, 21

Hu, W. and White, M. (1997). A CMB polarization primer. New A, 2:323–344. 22

Hubble, E. and Humason, M. L. (1931). The Velocity-Distance Relation among Extra-Galactic Nebulae. ApJ, 74:43. 1, 3

Hui, L. and Gnedin, N. Y. (1997). Equation of state of the photoionized intergalactic medium. *Monthly Notices of the Royal Astronomical Society*, 292(1):27–42. 57, 100, 105

Hummels, C. B. and Bryan, G. L. (2012). Adaptive Mesh Refinement Simulations of Galaxy Formation: Exploring Numerical and Physical Parameters. ApJ, 749:140. 109

Ikeuchi, S. (1986). The baryon clump within an extended dark matter region. Ap&SS, 118:509–514. 19

Iliev, I. T., Ciardi, B., Alvarez, M. A., Maselli, A., Ferrara, A., Gnedin, N. Y., Mellema, G., Nakamoto, T., Norman, M. L., Razoumov, A. O., Rijkhorst, E.-J., Ritzerveld, J., Shapiro, P. R., Susa, H., Umemura, M., and Whalen, D. J. (2006a). Cosmological radiative transfer codes comparison project - I. The static density field tests. MNRAS, 371:1057–1086. 32, 33, 39, 108, 109

Iliev, I. T., Mellema, G., Ahn, K., Shapiro, P. R., Mao, Y., and Pen, U.-L. (2014). Simulating cosmic reionization: how large a volume is large enough? MNRAS, 439:725–743. 19, 59

Iliev, I. T., Mellema, G., Pen, U.-L., Merz, H., Shapiro, P. R., and Alvarez, M. A. (2006b). Simulating cosmic reionization at large scales - I. The geometry of reionization. MNRAS, 369:1625–1638. 19, 59, 88, 92

Iliev, I. T., Mellema, G., Shapiro, P. R., and Pen, U.-L. (2007a). Self-regulated reionization. MNRAS, 376:534–548. 59, 84, 86, 88, 92

Iliev, I. T., Pen, U.-L., Bond, J. R., Mellema, G., and Shapiro, P. R. (2007b). The Kinetic Sunyaev-Zel'dovich Effect from Radiative Transfer Simulations of Patchy Reionization. ApJ, 660:933–944. 135

Iliev, I. T., Scannapieco, E., and Shapiro, P. R. (2005a). The Impact of Small-Scale Structure on Cosmological Ionization Fronts and Reionization. ApJ, 624:491–504. 55

Iliev, I. T., Shapiro, P. R., and Raga, A. C. (2005b). Minihalo photoevaporation during cosmic reionization: evaporation times and photon consumption rates. MNRAS, 361:405–414. 55, 61, 87

Iliev, I. T., Whalen, D., Mellema, G., Ahn, K., Baek, S., Gnedin, N. Y., Kravtsov, A. V., Norman, M., Raicevic, M., Reynolds, D. R., Sato, D., Shapiro, P. R., Semelin, B., Smidt, J., Susa, H., Theuns, T., and Umemura, M. (2009). Cosmological radiative transfer comparison project - II. The radiation-hydrodynamic tests. MNRAS, 400:1283–1316. 32, 33, 108, 109

Iwata, I., Inoue, A. K., Matsuda, Y., Furusawa, H., Hayashino, T., Kousai, K., Akiyama, M., Yamada, T., Burgarella, D., and Deharveng, J.-M. (2009). Detections of Lyman Continuum from Star-Forming Galaxies at z ˜ 3 through Subaru/Suprime-Cam Narrow-Band Imaging. ApJ, 692:1287–1293. 98

Katz, N., Quinn, T., Bertschinger, E., and Gelb, J. M. (1994). Formation of Quasars at High Redshift. MNRAS, 270:L71. 34

Khokhlov, A. (1998). Fully Threaded Tree Algorithms for Adaptive Refinement Fluid Dynamics Simulations. *Journal of Computational Physics*, 143:519–543. 40, 57, 100

Kim, J.-h., Krumholz, M. R., Wise, J. H., Turk, M. J., Goldbaum, N. J., and Abel, T. (2013). Dwarf Galaxies with Ionizing Radiation Feedback. I. Escape of Ionizing Photons. ApJ, 775:109. 98, 123

Kimm, T. and Cen, R. (2014). Escape Fraction of Ionizing Photons during Reionization: Effects due to Supernova Feedback and Runaway OB Stars. ApJ, 788:121. 44, 61, 98, 99, 105, 108, 109, 110, 121, 123, 158

Kimm, T., Katz, H., Haehnelt, M., Rosdahl, J., Devriendt, J., and Slyz, A. (2017). Feedback-regulated star formation and escape of LyC photons from mini-haloes during reionization. MNRAS, 466:4826–4846. 98, 99, 121, 122, 123

Kitsionas, S. and Whitworth, A. P. (2002). Smoothed Particle Hydrodynamics with particle splitting, applied to self-gravitating collapse. MNRAS, 330:129–136. 40

Komatsu, E., Smith, K. M., Dunkley, J., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., Nolta, M. R., Page, L., Spergel, D. N., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., Wollack, E., and Wright, E. L. (2011). Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. ApJS, 192:18. 58, 101, 144

Kravtsov, A. V., Klypin, A. A., and Khokhlov, A. M. (1997). Adaptive Refinement Tree: A New High-Resolution N-Body Code for Cosmological Simulations. ApJS, 111:73–94. 40

Kroupa, P. (2002). The Initial Mass Function of Stars: Evidence for Uniformity in Variable Systems. *Science*, 295:82–91. 28, 103

Kuhlen, M. and Faucher-Giguère, C.-A. (2012). Concordance models of reionization: implications for faint galaxies and escape fraction evolution. MNRAS, 423:862–876. 98, 99, 123

Lacey, C. G., Baugh, C. M., Frenk, C. S., Benson, A. J., Bower, R. G., Cole, S., Gonzalez-Perez, V., Helly, J. C., Lagos, C. D. P., and Mitchell, P. D. (2016). A unified multi-wavelength model of galaxy formation. MNRAS, 462:3854–3911. 74, 81

Larson, R. B. (2003). The physics of star formation. *Reports on Progress in Physics*, 66:1651–1697. 27

Leitet, E., Bergvall, N., Hayes, M., Linné, S., and Zackrisson, E. (2013). Escape of Lyman continuum radiation from local galaxies. Detection of leakage from the young starburst Tol 1247-232. A&A, 553:A106. 98, 123

Leitet, E., Bergvall, N., Piskunov, N., and Andersson, B.-G. (2011). Analyzing low signal-to-noise FUSE spectra. Confirmation of Lyman continuum escape from Haro 11. A&A, 532:A107. 98, 123

Leitherer, C., Ferguson, H. C., Heckman, T. M., and Lowenthal, J. D. (1995). The Lyman Continuum in Starburst Galaxies Observed with the Hopkins Ultraviolet Telescope. ApJ, 454:L19. 98, 113

Leitherer, C., Schaerer, D., Goldader, J. D., Delgado, R. M. G., Robert, C., Kune, D. F., de Mello, D. F., Devost, D., and Heckman, T. M. (1999). Starburst99: Synthesis Models for Galaxies with Active Star Formation. ApJS, 123:3–40. 29

Lewis, A., Challinor, A., and Lasenby, A. (2000). Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. ApJ, 538:473–476. xi, 15, 133

L'Huillier, B., Park, C., and Kim, J. (2014). Effects of the initial conditions on cosmological N-body simulations. New A, 30:79–88. 133

Lisenfeld, U. and Ferrara, A. (1998). Dust-to-Gas Ratio and Metal Abundance in Dwarf Galaxies. ApJ, 496:145–154. 108

Loeb, A. and Barkana, R. (2001). The Reionization of the Universe by the First Stars and Quasars. ARA&A, 39:19–66. 19

Loeb, A. and Furlanetto, S. R. (2013). *The First Galaxies in the Universe*. 17, 33, 122, 139

Ma, C.-P. and Bertschinger, E. (1995). Cosmological Perturbation Theory in the Synchronous and Conformal Newtonian Gauges. ApJ, 455:7. 17, 22

Ma, X., Hopkins, P. F., Kasen, D., Quataert, E., Faucher-Giguère, C.-A., Kereš, D., Murray, N., and Strom, A. (2016). Binary stars can provide the 'missing photons' needed for reionization. MNRAS, 459:3614–3619. 99, 101, 115, 121, 123, 124, 156

Ma, X., Kasen, D., Hopkins, P. F., Faucher-Giguère, C.-A., Quataert, E., Kereš, D., and Murray, N. (2015). The difficulty of getting high escape fractions of ionizing photons from high-redshift galaxies: a view from the FIRE cosmological simulations. MNRAS, 453:960–975. 98

Machacek, M. E., Bryan, G. L., and Abel, T. (2001). Simulations of Pregalactic Structure Formation with Radiative Feedback. ApJ, 548:509–521. 27

Madau, P., Haardt, F., and Rees, M. J. (1999). Radiative Transfer in a Clumpy Universe. III. The Nature of Cosmological Ionizing Sources. ApJ, 514:648–659. 97

Madau, P., Meiksin, A., and Rees, M. J. (1997). 21 Centimeter Tomography of the Intergalactic Medium at High Redshift. ApJ, 475:429–444. 128

Maio, U., Koopmans, L. V. E., and Ciardi, B. (2011). The impact of primordial supersonic flows on early structure formation, reionization and the lowest-mass dwarf galaxies. MNRAS, 412:L40–L44. 18, 127, 149, 152, 153

Marigo, P., Girardi, L., Bressan, A., Groenewegen, M. A. T., Silva, L., and Granato, G. L. (2008). Evolution of asymptotic giant branch stars. II. Optical to far-infrared isochrones with improved TP-AGB models. A&A, 482:883–905. 29

Martel, H. and Shapiro, P. R. (1998). A convenient set of comoving cosmological variables and their application. MNRAS, 297:467–485. 39

Martin, C. L. (2005). Mapping Large-Scale Gaseous Outflows in Ultraluminous Galaxies with Keck II ESI Spectra: Variations in Outflow Velocity with Galactic Mass. ApJ, 621:227–245. 44

Martin, C. L., Shapley, A. E., Coil, A. L., Kornei, K. A., Bundy, K., Weiner, B. J., Noeske, K. G., and Schiminovich, D. (2012). Demographics and Physical Properties of Gas Outflows/Inflows at $0.4 < z < 1.4$. ApJ, 760:127. 44

McGreer, I. D., Mesinger, A., and D'Odorico, V. (2015). Model-independent evidence in favour of an end to reionization by $z \approx 6$. MNRAS, 447:499–505. 63, 104

McGreer, I. D., Mesinger, A., and Fan, X. (2011). The first (nearly) model-independent constraint on the neutral hydrogen fraction at $z \sim 6$. MNRAS, 415:3237–3246. 22, 63, 104

McKee, C. F. and Ostriker, J. P. (1977). A theory of the interstellar medium - Three components regulated by supernova explosions in an inhomogeneous substrate. ApJ, 218:148–169. 44

McQuinn, M., Lidz, A., Zahn, O., Dutta, S., Hernquist, L., and Zaldarriaga, M. (2007). The morphology of HII regions during reionization. MNRAS, 377:1043–1063. 19, 59, 88, 92

McQuinn, M., Lidz, A., Zaldarriaga, M., Hernquist, L., and Dutta, S. (2008). Probing the neutral fraction of the IGM with GRBs during the epoch of reionization. MNRAS, 388:1101–1110. 63, 104

McQuinn, M. and O'Leary, R. M. (2012). The Impact of the Supersonic Baryon-Dark Matter Velocity Difference on the z ˜ 20 21 cm Background. ApJ, 760:3. 127, 128

Mellema, G., Iliev, I. T., Alvarez, M. A., and Shapiro, P. R. (2006a). C ^{2}-ray: A new method for photon-conserving transport of ionizing radiation. New A, 11:374–395. 32, 59

Mellema, G., Iliev, I. T., Pen, U.-L., and Shapiro, P. R. (2006b). Simulating cosmic reionization at large scales - II. The 21-cm emission features and statistical signals. MNRAS, 372:679–692. 23, 59, 88

Mesinger, A. and Furlanetto, S. (2007). Efficient Simulations of Early Structure Formation and Reionization. ApJ, 669:663–675. 59, 88

Meynet, G. and Maeder, A. (2000). Stellar evolution with rotation. V. Changes in all the outputs of massive star models. A&A, 361:101–120. 29

Mihalas, D. and Mihalas, B. W. (1984). *Foundations of radiation hydrodynamics*. 37

Mo, H., van den Bosch, F. C., and White, S. (2010). *Galaxy Formation and Evolution*. 16

Moore, B., Ghigna, S., Governato, F., Lake, G., Quinn, T., Stadel, J., and Tozzi, P. (1999). Dark Matter Substructure within Galactic Halos. ApJ, 524:L19–L22. 55

Moster, B. P., Naab, T., and White, S. D. M. (2013). Galactic star formation and accretion histories from matching galaxies to dark matter haloes. MNRAS, 428:3121–3138. xiii, xvi, 66, 70, 116

Nagashima, M., Gouda, N., and Sugiura, N. (1999). Effects of the UV background radiation on galaxy formation. MNRAS, 305:449–456. 55

Nakamoto, T., Umemura, M., and Susa, H. (2001). The effects of radiative transfer on the reionization of an inhomogeneous universe. MNRAS, 321:593–604. 92

Naoz, S. and Barkana, R. (2005). Growth of linear perturbations before the era of the first galaxies. MNRAS, 362:1047–1053. 129, 130, 153

Naoz, S. and Barkana, R. (2007). The formation and gas content of high-redshift galaxies and minihaloes. MNRAS, 377:667–676. 130, 153

Naoz, S., Barkana, R., and Mesinger, A. (2009). Gas in simulations of high-redshift galaxies and minihaloes. MNRAS, 399:369–376. 20

Naoz, S., Yoshida, N., and Gnedin, N. Y. (2012). Simulations of Early Baryonic Structure Formation with Stream Velocity. I. Halo Abundance. ApJ, 747:128. 18, 127, 152, 153

Naoz, S., Yoshida, N., and Gnedin, N. Y. (2013). Simulations of Early Baryonic Structure Formation with Stream Velocity. II. The Gas Fraction. ApJ, 763:27. 18, 127, 158

Netterfield, C. B., Ade, P. A. R., Bock, J. J., Bond, J. R., Borrill, J., Boscaleri, A., Coble, K., Contaldi, C. R., Crill, B. P., de Bernardis, P., Farese, P., Ganga, K., Giacometti, M., Hivon, E., Hristov, V. V., Iacoangeli, A., Jaffe, A. H., Jones, W. C., Lange, A. E., Martinis, L., Masi, S., Mason, P., Mauskopf, P. D., Melchiorri, A., Montroy, T., Pascale, E., Piacentini, F., Pogosyan, D., Pongetti, F., Prunet, S., Romeo, G., Ruhl, J. E., and Scaramuzzi, F. (2002). A Measurement by BOOMERANG of Multiple Peaks in the Angular Power Spectrum of the Cosmic Microwave Background. ApJ, 571:604–614. 5

Newman, S. F., Genzel, R., Förster-Schreiber, N. M., Shapiro Griffin, K., Mancini, C., Lilly, S. J., Renzini, A., Bouché, N., Burkert, A., Buschkamp, P., Carollo, C. M., Cresci, G., Davies, R., Eisenhauer, F., Genel, S., Hicks, E. K. S., Kurk, J., Lutz, D., Naab, T., Peng, Y., Sternberg, A., Tacconi, L. J., Vergani, D., Wuyts, S., and Zamorani, G. (2012). The SINS/zC-SINF Survey of z ˜ 2 Galaxy Kinematics: Outflow Properties. ApJ, 761:43. 44

Noh, Y. and McQuinn, M. (2014). A physical understanding of how reionization suppresses accretion on to dwarf haloes. MNRAS, 444:503–514. 20, 55, 56, 79

Ocvirk, P., Gillet, N., Aubert, D., Knebe, A., Libeskind, N., Chardin, J., Gottlöber, S., Yepes, G., and Hoffman, Y. (2014). The Reionization of Galactic Satellite Populations. ApJ, 794:20. 66

Ocvirk, P., Gillet, N., Shapiro, P. R., Aubert, D., Iliev, I. T., Teyssier, R., Yepes, G., Choi, J.-H., Sullivan, D., Knebe, A., Gottlöber, S., D'Aloisio, A., Park, H., Hoffman, Y., and Stranex, T. (2016). Cosmic Dawn (CoDa): the First Radiation-Hydrodynamics Simulation of Reionization and Galaxy Formation in the Local Universe. MNRAS, 463:1462–1485. 40, 61, 66, 86, 90, 92, 94

Okamoto, T., Gao, L., and Theuns, T. (2008). Mass loss of galaxies due to an ultraviolet background. MNRAS, 390:920–928. 20, 45, 55, 56, 61, 68, 79, 83, 84, 85, 87, 88, 89, 90, 95

O'Leary, R. M. and McQuinn, M. (2012). The Formation of the First Cosmic Structures and the Physics of the z ˜ 20 Universe. ApJ, 760:4. 18, 127, 128, 130, 133, 149, 152, 153

Ota, K., Iye, M., Kashikawa, N., Shimasaku, K., Kobayashi, M., Totani, T., Nagashima, M., Morokuma, T., Furusawa, H., Hattori, T., Matsuda, Y., Hashimoto, T., and Ouchi, M. (2008). Reionization and Galaxy Evolution Probed by z = 7 Lyα Emitters. ApJ, 677:12–26. 63, 104

Ouchi, M., Shimasaku, K., Furusawa, H., Saito, T., Yoshida, M., Akiyama, M., Ono, Y., Yamada, T., Ota, K., Kashikawa, N., Iye, M., Kodama, T., Okamura, S., Simpson, C., and Yoshida, M. (2010). Statistics of 207 Lyα Emitters at a Redshift Near 7: Constraints on Reionization and Galaxy Formation Models. ApJ, 723:869–894. 63, 104

Paardekooper, J.-P., Khochfar, S., and Dalla Vecchia, C. (2015). The First Billion Years project: the escape fraction of ionizing photons in the epoch of reionization. MNRAS, 451:2544–2563. 98, 99, 110

Paardekooper, J.-P., Pelupessy, F. I., Altay, G., and Kruip, C. J. H. (2011). The escape of ionising radiation from high-redshift dwarf galaxies. A&A, 530:A87. 98, 123

Padoan, P., Haugbølle, T., and Nordlund, Å. (2012). A Simple Law of Star Formation. ApJ, 759:L27. 47

Parsa, S., Dunlop, J. S., McLure, R. J., and Mortlock, A. (2016). The galaxy UV luminosity function at z $\sim$ 2-4; new results on faint-end slope and the evolution of luminosity density. MNRAS, 456:3194–3211. 115, 122

Pawlik, A. H. and Schaye, J. (2009). Photoheating and supernova feedback amplify each other's effect on the cosmic star formation rate. MNRAS, 396:L46–L50. 64, 67

Pawlik, A. H., Schaye, J., and Dalla Vecchia, C. (2015). Spatially adaptive radiation-hydrodynamical simulations of galaxy formation during cosmological reionization. MNRAS, 451:1586–1605. 92

Peacock, J. A. (1999). *Cosmological Physics*. 13

Peebles, P. J. E. (1980). *The large-scale structure of the universe*. 10

Peebles, P. J. E. and Yu, J. T. (1970). Primeval Adiabatic Perturbation in an Expanding Universe. ApJ, 162:815. 22

Penzias, A. A. and Wilson, R. W. (1965). A Measurement of Excess Antenna Temperature at 4080 Mc/s. ApJ, 142:419–421. 1

Perret, V. (2016). DICE: Disk Initial Conditions Environment. Astrophysics Source Code Library. xii, 47, 49, 50

Pietrinferni, A., Cassisi, S., Salaris, M., and Castelli, F. (2004). A Large Stellar Evolution Database for Population Synthesis Studies. I. Scaled Solar Models and Isochrones. ApJ, 612:168–190. 29

Planck Collaboration, Adam, R., Aghanim, N., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Battye, R., Benabed, K., Bernard, J.-P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bonaldi, A., Bonavera, L., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Calabrese, E., Cardoso, J.-F., Carron, J., Chiang, H. C., Colombo, L. P. L., Combet, C., Comis, B., Couchot, F., Coulais, A., Crill, B. P., Curto, A., Cuttaia, F., Davis, R. J., de Bernardis, P., de Rosa, A., de Zotti, G., Delabrouille, J., Di Valentino, E., Dickinson, C., Diego, J. M., Doré, O., Douspis, M., Ducout, A., Dupac, X., Elsner, F., Enßlin, T. A., Eriksen, H. K., Falgarone, E., Fantaye, Y., Finelli, F., Forastieri, F., Frailis, M., Fraisse, A. A., Franceschi, E., Frolov, A., Galeotta, S., Galli, S., Ganga, K., Génova-Santos, R. T., Gerbino, M., Ghosh, T., González-Nuevo, J., Górski, K. M., Gruppuso, A., Gudmundsson, J. E., Hansen, F. K., Helou, G., Henrot-Versillé, S., Herranz, D., Hivon, E., Huang, Z., Ilić, S., Jaffe, A. H., Jones, W. C., Keihänen, E., Keskitalo, R., Kisner, T. S., Knox, L., Krachmalnicoff, N., Kunz, M., Kurki-Suonio, H., Lagache, G., Lähteenmäki, A., Lamarre, J.-M., Langer, M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Le Jeune, M., Levrier, F., Lewis, A., Liguori, M., Lilje, P. B., López-Caniego, M., Ma, Y.-Z., Macías-Pérez, J. F., Maggio, G., Mangilli, A., Maris, M., Martin, P. G., Martínez-González, E., Matarrese, S., Mauri, N., McEwen, J. D., Meinhold, P. R., Melchiorri, A., Mennella, A., Migliaccio, M., Miville-Deschênes, M.-A., Molinari, D., Moneti, A., Montier, L., Morgante, G., Moss, A., Naselsky, P., Natoli, P., Oxborrow, C. A., Pagano, L., Paoletti, D., Partridge, B., Patanchon, G., Patrizii, L., Perdereau, O., Perotto, L., Pettorino, V., Piacentini, F., Plaszczynski, S., Polastri, L., Polenta, G., Puget, J.-L., Rachen, J. P., Racine, B., Reinecke, M., Remazeilles, M., Renzi, A., Rocha, G., Rossetti, M., Roudier, G., Rubiño-Martín, J. A., Ruiz-Granados, B., Salvati, L., Sandri, M., Savelainen, M., Scott, D., Sirri, G., Sunyaev, R., Suur-Uski, A.-S., Tauber, J. A., Tenti, M., Toffolatti, L., Tomasi, M., Tristram, M.,

Trombetti, T., Valiviita, J., Van Tent, F., Vielva, P., Villa, F., Vittorio, N., Wandelt, B. D., Wehus, I. K., White, M., Zacchei, A., and Zonca, A. (2016a). Planck intermediate results. XLVII. Planck constraints on reionization history. A&A, 596:A108. x, 6, 7, 21, 23, 60, 63, 90, 103, 104

Planck Collaboration, Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Baker, M., Balbi, A., Banday, A. J., and et al. (2011). Planck early results. I. The Planck mission. A&A, 536:A1. 5

Planck Collaboration, Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., Bartlett, J. G., and et al. (2016b). Planck 2015 results. XIII. Cosmological parameters. A&A, 594:A13. x, 6, 7, 60, 103, 104

Poincare, H. (1906). The Milky Way and the Theory of Gases. *Popular Astronomy*, 14:475–488. 2

Pontzen, A., Roškar, R., Stinson, G., and Woods, R. (2013). pynbody: N-Body/SPH analysis for python. Astrophysics Source Code Library. 91, 125, 154

Press, W. H. and Schechter, P. (1974). Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. ApJ, 187:425–438. 149

Pritchard, J. R. and Loeb, A. (2012). 21 cm cosmology in the 21st century. *Reports on Progress in Physics*, 75(8):086901. 23

Quinn, T., Katz, N., and Efstathiou, G. (1996). Photoionization and the formation of dwarf galaxies. MNRAS, 278:L49–L54. 20, 55, 84

Rasera, Y. and Teyssier, R. (2006). The history of the baryon budget. Cosmic logistics in a hierarchical universe. A&A, 445:1–27. 42

Razoumov, A. O. and Sommer-Larsen, J. (2006). Escape of Ionizing Radiation from Star-forming Regions in Young Galaxies. ApJ, 651:L89–L92. 92, 98, 123

Razoumov, A. O. and Sommer-Larsen, J. (2007). Modeling Lyman Continuum Emission from Young Galaxies. ApJ, 668:674–681. 98, 123

Razoumov, A. O. and Sommer-Larsen, J. (2010). Ionizing Radiation from z = 4-10 Galaxies. ApJ, 710:1239–1246. 98

Rees, M. J. (1986). Lyman absorption lines in quasar spectra - Evidence for gravitationally-confined gas in dark minihaloes. MNRAS, 218:25P–30P. 19

Rees, M. J. (1999). The end of the 'dark age'. In Holt, S. and Smith, E., editors, *After the Dark Ages: When Galaxies were Young (the Universe at 2 < Z < 5)*, volume 470 of *American Institute of Physics Conference Series*, pages 13–23. 1, 54

Ricotti, M., Gnedin, N. Y., and Shull, J. M. (2002). The Fate of the First Galaxies. I. Self-consistent Cosmological Simulations with Radiative Transfer. ApJ, 575:33–48. 19

Robertson, B. E., Ellis, R. S., Furlanetto, S. R., and Dunlop, J. S. (2015). Cosmic Reionization and Early Star-forming Galaxies: A Joint Analysis of New Constraints from Planck and the Hubble Space Telescope. ApJ, 802:L19. 63, 99, 104, 123

Robertson, B. E., Furlanetto, S. R., Schneider, E., Charlot, S., Ellis, R. S., Stark, D. P., McLure, R. J., Dunlop, J. S., Koekemoer, A., Schenker, M. A., Ouchi, M., Ono, Y., Curtis-Lake, E., Rogers, A. B., Bowler, R. A. A., and Cirasuolo, M. (2013). New Constraints on Cosmic Reionization from the 2012 Hubble Ultra Deep Field Campaign. ApJ, 768:71. 98, 99, 123

Rosdahl, J. and Blaizot, J. (2012). Extended Lyα emission from cold accretion streams. MNRAS, 423:344–366. 87, 92

Rosdahl, J., Blaizot, J., Aubert, D., Stranex, T., and Teyssier, R. (2013). RAMSES-RT: radiation hydrodynamics in the cosmological context. MNRAS, 436:2188–2231. 31, 33, 34, 37, 39, 45, 57, 89, 100, 108, 109

Rosdahl, J., Schaye, J., Teyssier, R., and Agertz, O. (2015). Galaxies that shine: radiation-hydrodynamical simulations of disc galaxies. MNRAS, 451:34–58. 42, 61, 86

Ross, H. E., Dixon, K. L., Iliev, I. T., and Mellema, G. (2017). Simulating the impact of X-ray heating during the cosmic dawn. MNRAS, 468:3785–3797. 24

Roškar, R., Teyssier, R., Agertz, O., Wetzstein, M., and Moore, B. (2014). A systematic look at the effects of radiative feedback on disc galaxy formation. MNRAS, 444:2837–2853. 44, 55

Rupke, D. S., Veilleux, S., and Sanders, D. B. (2005). Outflows in Infrared-Luminous Starbursts at z < 0.5. II. Analysis and Discussion. ApJS, 160:115–148. 44

Schaller, G., Schaerer, D., Meynet, G., and Maeder, A. (1992). New grids of stellar models from 0.8 to 120 solar masses at Z = 0.020 and Z = 0.001. A&AS, 96:269–331. 29

Schroeder, J., Mesinger, A., and Haiman, Z. (2013). Evidence of Gunn-Peterson damping wings in high-z quasar spectra: strengthening the case for incomplete reionization at z ∼ 6-7. MNRAS, 428:3058–3071. 63, 104

Semelin, B. (2016). Detailed modelling of the 21-cm forest. MNRAS, 455:962–973. 24

Semelin, B., Combes, F., and Baek, S. (2007). Lyman-alpha radiative transfer during the epoch of reionization: contribution to 21-cm signal fluctuations. A&A, 474:365–374. 24

Shapiro, P. R., Giroux, M. L., and Babul, A. (1994). Reionization in a cold dark matter universe: The feedback of galaxy formation on the intergalactic medium. ApJ, 427:25–50. 20, 55, 56, 61, 84

Shapiro, P. R., Iliev, I. T., Alvarez, M. A., and Scannapieco, E. (2006). Relativistic Ionization Fronts. ApJ, 648:922–935. 89

Shapiro, P. R., Iliev, I. T., and Raga, A. C. (2004). Photoevaporation of cosmological minihaloes during reionization. MNRAS, 348:753–782. 33, 55, 61, 87

Sheth, R. K. and Tormen, G. (2002). An excursion set model of hierarchical clustering: ellipsoidal collapse and the moving barrier. MNRAS, 329:61–75. 149, 150

Shimabukuro, H. and Semelin, B. (2017). Analysing the 21 cm signal from the epoch of reionization with artificial neural networks. MNRAS, 468:3869–3877. 24, 75

Shull, J. M., Harness, A., Trenti, M., and Smith, B. D. (2012). Critical Star Formation Rates for Reionization: Full Reionization Occurs at Redshift z ≈ 7. ApJ, 747:100. 98

Siana, B., Shapley, A. E., Kulas, K. R., Nestor, D. B., Steidel, C. C., Teplitz, H. I., Alavi, A., Brown, T. M., Conselice, C. J., Ferguson, H. C., Dickinson, M., Giavalisco, M., Colbert, J. W., Bridge, C. R., Gardner, J. P., and de Mello, D. F. (2015). A Deep Hubble Space Telescope and Keck Search for Definitive Identification of Lyman Continuum Emitters at z˜3.1. ApJ, 804:17. 98

Smoot, G. F., Bennett, C. L., Kogut, A., Wright, E. L., Aymon, J., Boggess, N. W., Cheng, E. S., de Amici, G., Gulkis, S., Hauser, M. G., Hinshaw, G., Jackson, P. D., Janssen, M., Kaita, E., Kelsall, T., Keegstra, P., Lineweaver, C., Loewenstein, K., Lubin, P., Mather, J., Meyer, S. S., Moseley, S. H., Murdock, T., Rokke, L., Silverberg, R. F.,

Tenorio, L., Weiss, R., and Wilkinson, D. T. (1992). Structure in the COBE differential microwave radiometer first-year maps. ApJ, 396:L1–L5. 5

Somerville, R. S. (2002). Can Photoionization Squelching Resolve the Substructure Crisis? ApJ, 572:L23–L26. 55

Somerville, R. S., Bullock, J. S., and Livio, M. (2003). The Epoch of Reionization in Models with Reduced Small-Scale Power. ApJ, 593:616–621. 98

Springel, V. (2005). The cosmological simulation code GADGET-2. MNRAS, 364:1105–1134. 40, 88

Springel, V. and Hernquist, L. (2003). Cosmological smoothed particle hydrodynamics simulations: a hybrid multiphase model for star formation. MNRAS, 339:289–311. 44

Srisawat, C. (2016). *Semi-analytic model of galaxy formation with radiative feedback during the Epoch of Reionisation*. PhD thesis, University of Sussex. 90

Stacy, A., Bromm, V., and Loeb, A. (2011). Effect of Streaming Motion of Baryons Relative to Dark Matter on the Formation of the First Stars. ApJ, 730:L1. 18, 127, 149, 152, 153

Steidel, C. C., Rudie, G. C., Strom, A. L., Pettini, M., Reddy, N. A., Shapley, A. E., Trainor, R. F., Erb, D. K., Turner, M. L., Konidaris, N. P., Kulas, K. R., Mace, G., Matthews, K., and McLean, I. S. (2014). Strong Nebular Line Ratios in the Spectra of z ˜ 2-3 Star Forming Galaxies: First Results from KBSS-MOSFIRE. ApJ, 795:165. 101

Sullivan, D., Iliev, I. T., and Dixon, K. L. (2017). Using Artificial Neural Networks to Constrain the Halo Baryon Fraction during Reionization. *ArXiv e-prints*. 54

Sunyaev, R. A. and Zeldovich, Y. B. (1972). The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies. *Comments on Astrophysics and Space Physics*, 4:173. 22

Teyssier, R. (2002). Cosmological hydrodynamics with adaptive mesh refinement. A new high resolution code called RAMSES. A&A, 385:337–364. 34, 36, 40, 57, 100

Thornton, K., Gaudlitz, M., Janka, H.-T., and Steinmetz, M. (1998). Energy Input and Mass Redistribution by Supernovae in the Interstellar Medium. ApJ, 500:95–119. 110

Toro, E. F., Spruce, M., and Speares, W. (1994). Restoration of the contact surface in the HLL-Riemann solver. *Shock Waves*, 4:25–34. 57, 100

Trac, H. and Cen, R. (2007). Radiative Transfer Simulations of Cosmic Reionization. I. Methodology and Initial Results. ApJ, 671:1–13. 19

Trebitsch, M., Blaizot, J., Rosdahl, J., Devriendt, J., and Slyz, A. (2017). Fluctuating feedback-regulated escape fraction of ionizing radiation in low-mass, high-redshift galaxies. MNRAS, 470:224–239. 98, 99, 108, 121, 123

Truelove, J. K., Klein, R. I., McKee, C. F., Holliman, II, J. H., Howell, L. H., and Greenough, J. A. (1997). The Jeans Condition: A New Constraint on Spatial Resolution in Simulations of Isothermal Self-gravitational Hydrodynamics. ApJ, 489:L179–L183. 42

Tseliakhovich, D., Barkana, R., and Hirata, C. M. (2011). Suppression and spatial variation of early galaxies and minihaloes. MNRAS, 418:906–915. xi, 13, 14, 127

Tseliakhovich, D. and Hirata, C. (2010). Relative velocity of dark matter and baryonic fluids and the formation of the first structures. Phys. Rev. D, 82(8):083520. xvii, 17, 18, 126, 127, 133, 139, 141, 153

Umeda, H. and Nomoto, K. (2003). First-generation black-hole-forming supernovae and the metal abundance pattern of a very iron-poor star. Nature, 422:871–873. 27

Vacondio, R., Rogers, B. D., Stansby, P. K., Mignosa, P., and Feldman, J. (2013). Variable resolution for SPH: A dynamic particle coalescing and splitting scheme. *Computer Methods in Applied Mechanics and Engineering*, 256:132–148. 40

Vogelsberger, M., Genel, S., Springel, V., Torrey, P., Sijacki, D., Xu, D., Snyder, G., Nelson, D., and Hernquist, L. (2014). Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe. MNRAS, 444:1518–1547. 34, 40, 144

Weinberg, D. H., Hernquist, L., and Katz, N. (1997). Photoionization, Numerical Resolution, and Galaxy Formation. ApJ, 477:8–20. 20, 55

Weiner, B. J., Coil, A. L., Prochaska, J. X., Newman, J. A., Cooper, M. C., Bundy, K., Conselice, C. J., Dutton, A. A., Faber, S. M., Koo, D. C., Lotz, J. M., Rieke, G. H., and Rubin, K. H. R. (2009). Ubiquitous Outflows in DEEP2 Spectra of Star-Forming Galaxies at z = 1.4. ApJ, 692:187–211. 44

Weinmann, S. M., Macciò, A. V., Iliev, I. T., Mellema, G., and Moore, B. (2007). Dependence of the local reionization history on halo mass and environment: did Virgo reionize the Local Group? MNRAS, 381:367–376. 66

White, S. D. M. and Rees, M. J. (1978). Core condensation in heavy halos - A two-stage theory for galaxy formation and clustering. MNRAS, 183:341–358. 24

Willott, C. J., Delorme, P., Reylé, C., Albert, L., Bergeron, J., Crampton, D., Delfosse, X., Forveille, T., Hutchings, J. B., McLure, R. J., Omont, A., and Schade, D. (2010). The Canada-France High-z Quasar Survey: Nine New Quasars and the Luminosity Function at Redshift 6. AJ, 139:906–918. 98

Wise, J. H. and Abel, T. (2011). ENZO+MORAY: radiation hydrodynamics adaptive mesh refinement simulations with adaptive ray tracing. MNRAS, 414:3458–3491. 31

Wise, J. H. and Cen, R. (2009). Ionizing Photon Escape Fractions From High-Redshift Dwarf Galaxies. ApJ, 693:984–999. 98

Wise, J. H., Demchenko, V. G., Halicek, M. T., Norman, M. L., Turk, M. J., Abel, T., and Smith, B. D. (2014). The birth of a galaxy - III. Propelling reionization with the faintest galaxies. MNRAS, 442:2560–2579. 98, 110, 149

Worthey, G. (1994). Comprehensive stellar population models and the disentanglement of age and metallicity effects. ApJS, 95:107–149. 29

Wyithe, J. S. B. and Bolton, J. S. (2011). Near-zone sizes and the rest-frame extreme ultraviolet spectral index of the highest redshift quasars. MNRAS, 412:1926–1936. 61, 63, 104

Wyithe, J. S. B. and Cen, R. (2007). The Extended Star Formation History of the First Generation of Stars and the Reionization of Cosmic Hydrogen. ApJ, 659:890–907. 98

Wyithe, J. S. B. and Loeb, A. (2006). Suppression of dwarf galaxy formation by cosmic reionization. Nature, 441:322–324. 61

Yoshida, N., Abel, T., Hernquist, L., and Sugiyama, N. (2003). Simulations of Early Structure Formation: Primordial Gas Clouds. ApJ, 592:645–663. 27

Zahn, O., Lidz, A., McQuinn, M., Dutta, S., Hernquist, L., Zaldarriaga, M., and Furlanetto, S. R. (2007). Simulations and Analytic Calculations of Bubble Growth during Hydrogen Reionization. ApJ, 654:12–26. 59, 88

Zahn, O., Mesinger, A., McQuinn, M., Trac, H., Cen, R., and Hernquist, L. E. (2011). Comparison of reionization models: radiative transfer simulations and approximate, seminumeric models. MNRAS, 414:727–738. 33

Zaldarriaga, M. (1997). Polarization of the microwave background in reionized models. Phys. Rev. D, 55:1822–1829. 21

Zaroubi, S. (2013). The Epoch of Reionization. In Wiklind, T., Mobasher, B., and Bromm, V., editors, *The First Galaxies*, volume 396 of *Astrophysics and Space Science Library*, page 45. 21, 22

Zastrow, J., Oey, M. S., Veilleux, S., McDonald, M., and Martin, C. L. (2011). An Ionization Cone in the Dwarf Starburst Galaxy NGC 5253. ApJ, 741:L17. 98, 113

Zwicky, F. (1933). Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127. 2